



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 124533

TO: Bao-Qun Li
Location: REM-3D24&3C18
Art Unit: 1648
Thursday, June 17, 2004

Case Serial Number: 09/899303

From: Toby Port
Location: Biotech-Chem Library
Remsen 1A59
Phone: 571-272-2523

toby.port@uspto.gov

Search Notes

Dear Examiner Li,

Here are the results of your search.
Please feel free to contact me if you have any questions.

Toby Port

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2004, 08:32:08 ; Search time 1958.63 Seconds
(without alignments)
14206.949 Million cell updates/sec

Title: US-09-899-303A-3

Perfect score: 642

Sequence: 1 ATGCCCGTGTCTCTTCTC.....TACTCTTGTCTCTTAATAG 642

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_hgt.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sta.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pin.*
- 35: em_htg_rod.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_hgo_hum.*
- 40: em_hgo_mus.*
- 41: em_hgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	642	100.0	642	6	A48665	A48665 Sequence 3
2	642	100.0	642	6	AR157324	AR157324 Sequence
3	642	100.0	642	6	AX452752	AX452752 Sequence
4	642	100.0	642	6	AX685004	AX685004 Sequence
5	642	100.0	642	6	AX794846	AX794846 Sequence
6	628.2	97.9	795	6	A48667	A48667 Sequence 5
7	628.2	97.9	795	6	AR157325	AR157325 Sequence
8	628.2	97.9	795	6	AX452754	AX452754 Sequence
9	628.2	97.9	795	6	AX685006	AX685006 Sequence
10	628.2	97.9	795	6	AX794848	AX794848 Sequence
11	619.8	96.5	2082	6	A48709	A48709 Sequence 47
12	619.8	96.5	2082	6	AR157350	AR157350 Sequence
13	619.8	96.5	2082	6	AX452796	AX452796 Sequence
14	619.8	96.5	2082	6	AX685048	AX685048 Sequence
15	619.8	96.5	2082	6	AX794890	AX794890 Sequence
16	619.8	96.5	2433	6	A48711	A48711 Sequence 49
17	619.8	96.5	2433	6	AR157351	AR157351 Sequence
18	619.8	96.5	2433	6	AX452798	AX452798 Sequence
19	619.8	96.5	2433	6	AX685050	AX685050 Sequence
20	619.8	96.5	2433	6	AX794892	AX794892 Sequence
21	575	89.6	9379	14	AF207766	AF207766 Hepatitis
22	573.4	89.3	9410	14	HPCXR2	D50481 Hepatitis C
23	569.6	88.7	9418	14	HCV132996	AU132996 Hepatitis
24	568.6	88.6	9386	14	AF165056	AF165056 Hepatitis
25	567	88.3	9386	14	AF165055	AF165055 Hepatitis
26	567	88.3	9410	14	HPCXR2	D50485 Hepatitis C
27	567	88.3	9598	14	AB049101	AB049101 Hepatitis
28	566.4	88.2	3461	6	E07544	E07544 Gene fragme
29	566.4	88.2	3461	6	E09288	E09288 cDNA encodi
30	565.4	88.1	1880	14	HPCSTRJ4	D00832 Hepatitis C
31	565.4	88.1	2540	6	E04260	E04260 cDNA encodi
32	565.4	88.1	2540	6	E04805	E04805 cDNA to 5'
33	565.4	88.1	2540	6	E07391	E07391 cDNA encodi
34	565.4	88.1	3296	14	AB008441	AB008441 Hepatitis
35	565.4	88.1	9033	14	HCU238800	AU238800 Hepatitis
36	565.4	88.1	9448	14	HPCJ483	D13558 Hepatitis C
37	565.4	88.1	9456	14	HPCRNA	D10934 Hepatitis C
38	565.4	88.1	9605	6	AX739971	AX739971 Sequence
39	565.4	88.1	9605	14	HCU238799	AJ238799 Hepatitis
40	565.4	88.1	11076	6	AR406040	AR406040 Sequence
41	565.4	88.1	11076	6	AR406046	AR406046 Sequence
42	565.4	88.1	11076	6	AR406048	AR406048 Sequence
43	565.4	88.1	11076	6	AR406050	AR406050 Sequence
44	565.4	88.1	11076	6	AX036252	AX036252 Sequence
45	565.4	88.1	11076	6	AX036258	AX036258 Sequence

ALIGNMENTS

RESULT 1

A48665	A48665	Sequence 3 from Patent WO9604385.	642 bp	DNA	linear	PAT 07-MAR-1997
LOCUS	Sequence 3 from Patent WO9604385.					
DEFINITION	A48665					
ACCESSION	A48665					
VERSION	A48665.1	GI:2302378				
KEYWORDS	unidentified					
SOURCE	unclassified					
ORGANISM	1 (bases 1 to 642)					
REFERENCE	Maertens,G., Bosman,F., De,M.G. and Buyse,M.					
AUTHORS	PURIFIED HEPATITIS C VIRUS ENVELOPE PROTEINS FOR DIAGNOSTIC AND					
TITLE	THERAPEUTIC USE					
JOURNAL	Patent: WO 9604385-A 3 15-FEB-1996;					

ORGANISM Hepatitis C virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.

REFERENCE 1
AUTHORS Maertens, G., Depla, E. and Bosman, F.
TITLE Purified Hepatitis C virus envelope proteins for diagnostic and
therapeutic use
JOURNAL Patent: WO 03051912-A 5 26-JUN-2003;
INNOGENETICS N.V. (BE)
FEATURES Location/Qualifiers
source 1..795
/organism="Hepatitis C virus"
/mol_type="unassigned DNA"
/db_xref="taxon:11103"
CDS 1..792
/note="unnamed protein product"
/codon_start=1
/protein_id="CA848243.1"
/db_xref="GI:37515738"
/translation="MLGKVIDTLTCGPADLVGYIPLVGAPLGGARALAHGVRVLEDG
VNYATGNLPGCSFSIFLLALLSCLTVPASAYEVNVSVMYHVTNDSCNSIVYBAADM
IMHTPGCVPCVRENSSRCWVALLPTLAARNASVPTTIRRHVDLLVGAARFCSAMTV
GDLGGSFVLVSQFLTISPRRHETVQDCNCSYPGHITGRMAMDMNMMNSPTTALVVS
QLLRIPQAVDMVAGAHGVLAYYSYMGVNAKVLIVMLLPAP"
mat_peptide 1..789
/product="unnamed"

ORIGIN
Query Match 97.9%; Score 628.2; DB 6; Length 795;
Best Local Similarity 98.8%; Pred. No. 4e-147;
Matches 633; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 2 TGCCCGGTGCTCTTCTCTATCTCTCTTGGCTTACTGTCTGTGACCAATTCAG 61
DB 155 TGCCCGGTGCTCTTCTCTATCTCTCTTGGCTTGGCTGTCTGTGACCGTTCAG 214
QY 62 CTTCCGCTTATGAGTCGCGCAACGTGTCCGGATGTACCATGTACGACGACGTCTCCA 121
DB 215 CTTCCGCTTATGAGTCGCGCAACGTGTCCGGATGTACCATGTACGACGACGTCTCCA 274
QY 122 ACTCAGCATTTGTATGAGCGACGACATGATCATGCACACCCCGGGTGCGCCT 181
DB 275 ACTCAGCATTTGTATGAGCGACGACATGATCATGCACACCCCGGGTGCGCCT 334
QY 182 GCCTTCGGGAGAACAACTCTTCCGCTGCTGGGTAGCGCTCACCCCGACGCTCCAGCTA 241
DB 335 GCCTTCGGGAGAACAACTCTTCCGCTGCTGGGTAGCGCTCACCCCGACGCTCCAGCTA 394
QY 242 GGAACGCGACGCTCCCGACCAATACGACCGACGCTGATTTGCTGTGGGCGG 301
DB 395 GGAACGCGACGCTCCCGACCAATACGACCGACGCTGATTTGCTGTGGGCGG 454
QY 302 CTGCTCTCTCTTCCGCTATGATGAGGGGATCTCTGGGATCTGTCTCCGCTCTCC 361
DB 455 CTGCTCTCTCTTCCGCTATGATGAGGGGACCTCTCGGATCTGTCTTCCGCTCTCC 514
QY 362 AGCTGTTTACCATCTCGCTCTCGCGCATGAGAGCGTGCAGACTGCAATTTGCTCAATCT 421
DB 515 AGCTGTTTACCATCTCGCTCTCGCGCATGAGAGCGTGCAGACTGCAATTTGCTCAATCT 574
QY 422 ATCCCGGCCACATAACAGGTACCGTATGCTTGGGATATGATGATGATGATGATGATGAT 481
DB 575 ATCCCGGCCACATAACAGGTACCGTATGCTTGGGATATGATGATGATGATGATGATGAT 634
QY 482 CAACGGCCCTTGGTGTATCGGAGTCTCGGATTCACCAAGCTGTGTGACATGTGTGG 541
DB 635 CAACGGCCCTTGGTGTATCGGAGTCTCGGATTCACCAAGCTGTGTGACATGTGTGG 694
QY 542 CGGGGGCCCATTTGGGAGTCTCTGCGGCCCTCGGCTACTATTTCATGTTGGGGAACGTGG 601
DB 695 CGGGGGCCCATTTGGGAGTCTCTGCGGGTCTCGGCTACTATTTCATGTTGGGGAACGTGG 754
QY 602 CTAAGGTTTTGATGTGATGCTACTCTTTGCTCTCTAATAG 642

Db 755 CTAAGGTTTTGATGTGATGCTACTCTTTGCTCCTAATAG 795

RESULT 11
LOCUS A48709 2082 bp DNA linear PAT 07-MAR-1997
DEFINITION Sequence 47 from Patent WO9604385.
ACCESSION A48709
VERSION A48709.1 GI:2302422
KEYWORDS
SOURCE unidentified
ORGANISM unidentified
unclassified.
REFERENCE 1 (bases 1 to 2082)
AUTHORS Maertens, G., Bosman, F., De, M. G. and Buyse, M.
TITLE PURIFIED HEPATITIS C VIRUS ENVELOPE PROTEINS FOR DIAGNOSTIC AND
THERAPEUTIC USE
JOURNAL Patent: WO 9604385-A 47 15-FEB-1996;
INNOGENETICS NV (BE)
COMMENT Other publication CA 2172273 960215
Other publication AU 3382495 960304.
FEATURES Location/Qualifiers
source 1..2082
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"
CDS 1..2079
/note="unnamed protein product; Protein sequence is in
conflict with the conceptual translation"
/codon_start=1
/protein_id="CAA03146.1"
/db_xref="GI:2302423"
/db_xref="REMTREMBL:CAA03146"
/translation="NLGKVIDTLTCGPADLVGYIPLVGAPLGGARALAHGVRVLEDG
VNYATGNLPGCSFSIFLLALLSCLTVPASAYEVNVSVMYHVTNDSCNSIVYBAADM
IMHTPGCVPCVRENSSRCWVALLPTLAARNASVPTTIRRHVDLLVGAARFCSAMTV
GDLGGSFVLVSQFLTISPRRHETVQDCNCSYPGHITGRMAMDMNMMNSPTTALVVS
QLLRIPQAVDMVAGAHGVLAYYSYMGVNAKVLIVMLLPAGDGHTRVSGGAA
SDTRFVLSLPGSAQKIQLVNTGNSWHINRTALNCNDSLOTGFFAALPYKHKFNSSG
CPRLASCRSIDKEAQGGLTYTSPNSDORPCYCHVAPRPGIVPASAOCGPVCP
TPSPVVTGTRDRCVPTYNWANDSDVILLNTRPPRGNWFGCTTMCNCTGPTKCGGP
PCNIGAGNNLTCTCTCFRKHPRATYARCGSGFWLTPRCMWHYFRLWHHPCTVNPFT
IFKRVYGVGVEHREFAACNWRGCDLDRSELSPLLLSLTTEDARICACIMMLLI
ALSTGLHLHONIVDQVLYGSAVSLVKWEYVLLFLLLADARICACIMMLLI
AOAEALLENLVLAANAAGVAGHGLTSLFVFPFAAWYIKGRVLPGAAYAFYGVPLLL
LLALPPRAYA"
mat_peptide 1..2076
/product="unnamed"

ORIGIN
Query Match 96.5%; Score 619.8; DB 6; Length 2082;
Best Local Similarity 98.9%; Pred. No. 5.3e-145;
Matches 624; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 2 TGCCCGGTGCTCTTCTCTATCTTCTTGGCTTTATCTGTCTGTGACCAATTCAG 61
DB 155 TGCCCGGTGCTCTTCTCTATCTTCTTGGCTTTGTCTGTCTGTGACCGTTCAG 214
QY 62 CTTCCGCTTATGAGTCGCGCAACGTGTCCGGATGTACCATGTACGACGACGTCTCCA 121
DB 215 CTTCCGCTTATGAGTCGCGCAACGTGTCCGGATGTACCATGTACGACGACGTCTCCA 274
QY 122 ACTCAGCATTTGTATGAGCGACGACATGATCATGCACACCCCGGGTGCGCCT 181
DB 275 ACTCAGCATTTGTATGAGCGACGACATGATCATGCACACCCCGGGTGCGCCT 334
QY 182 CGCTTCGGGAGAACAACTCTTCCGCTGCTGGGTAGCGCTCACCCCGACGCTCCAGCTA 241
DB 335 CGCTTCGGGAGAACAACTCTTCCGCTGCTGGGTAGCGCTCACCCCGACGCTCCAGCTA 394
QY 242 GGAACGCGACGCTCCCGACCAATACGACCGACGCTGATTTGCTGTGGGCGG 301

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2004, 07:39:53 ; Search time 241.553 Seconds
(without alignments)
11290.892 Million cell updates/sec

Title: US-09-899-303A-3

Perfect score: 642

Sequence: 1 ATGCCCGTGGCTCTCTC.....TACTCTTTGGCTCTCTAATAG 642

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_29Jan04.*

1: Geneseqn1980s.*

2: Geneseqn1990s.*

3: Geneseqn2000s.*

4: Geneseqn2001as.*

5: Geneseqn2001bs.*

6: Geneseqn2002as.*

7: Geneseqn2003as.*

8: Geneseqn2003bs.*

9: Geneseqn2003cs.*

10: Geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	642	100.0	642	2	AAT12704 HCV E1 co
2	642	100.0	642	6	AAL48913 Hepatitis
3	642	100.0	642	9	ADD55634 Hepatitis
4	628.2	97.9	795	2	AAT12705 HCV E1 co
5	628.2	97.9	795	6	AAL48914 Hepatitis
6	628.2	97.9	795	9	ADD55513 Hepatitis
7	619.8	96.5	2082	6	AAL48939 Hepatitis
8	619.8	96.5	2082	9	ADD55555 Hepatitis
9	619.8	96.5	2086	2	AAT12973 HCV E1 co
10	619.8	96.5	2433	2	AAT12974 HCV E1 co
11	619.8	96.5	2433	9	ADD55557 Hepatitis
12	608.8	94.8	2434	6	AAL48940 Hepatitis
13	566.4	88.2	3461	2	AAG64068 Non-A, no
14	566.4	88.2	3461	2	AAT30386 5'UTR/COR
15	565.4	88.1	1880	2	AAG24467 NANB hepa
16	565.4	88.1	2187	2	ABA03491 Cuticle p
17	565.4	88.1	2540	2	AAG43889 NANB hepa
18	565.4	88.1	2540	2	AAQ63753 NANBHV ge
19	565.4	88.1	9605	6	ABK91431 Hepatitis
20	565.4	88.1	9605	6	ABK91424 Hepatitis
21	565.4	88.1	9605	6	ABK91429 Hepatitis
22	565.4	88.1	9605	6	ABK91432 Hepatitis
23	565.4	88.1	9605	6	ABK91411 Hepatitis

ALIGNMENTS

RESULT 1
AAT12704

ID AAT12704 standard; DNA; 642 BP.

AC AAT12704;

DT 23-SEP-1996 (first entry)

XX HCV E1 construct HCC19A.

XX HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human;
KW serotype; reversed phase hybridisation assay; genotype; antigen; sera;
XX ss.

OS Hepatitis C virus.

XX WO9604385-A2.

XX 15-FEB-1996.

XX 31-JUL-1995; 95WO-EP003031.

XX 29-JUL-1994; 94EP-00870132.

XX (INNO-) INNOGENETICS NV.

XX Maertens G, Bosman F, De Martynoff G, Buysse M;

XX WPI; 1996-129401/13.

XX Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope proteins
PT - in presence of disulphide bond cleavage agent, to produce proteins
PT suitable for direct use in vaccines or diagnostic assays of HCV.

XX Claim 23; Fig 21; 146pp; English.

XX AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1
CC and E2 protein coding sequence constructs. These sequences are included
CC in vectors for the production of recombinant E1, E2, and E1/E2 proteins.
CC The recombinant proteins can then be isolated using a method of the
CC invention. In the method, the envelope proteins are purified by carrying
CC out a disulphide bond cleavage, or a reduction step with a disulphide
CC bond cleavage agent, after lysis of recombinant host cells. The
CC constructs containing the purified HCV envelope proteins can be used for
CC vaccinating humans against HCV, for in vitro detection of HCV antibodies
CC in a sample, and in a serotyping assay for detecting one or more

24 565.4 88.1 9605 6 ABK91430 Hepatitis
25 565.4 88.1 9605 6 ABK91428 Hepatitis
26 565.4 88.1 9605 6 ABK91425 Hepatitis
27 565.4 88.1 9605 6 ABK91426 Hepatitis
28 565.4 88.1 9605 6 ABK91433 Hepatitis
29 565.4 88.1 9605 6 AAD25332 Hepatitis
30 565.4 88.1 9608 6 ABK91427 Hepatitis
31 565.4 88.1 11062 6 AAD25331 Hepatitis
32 565.4 88.1 11076 3 AAQ98965 Hepatitis
33 563.8 87.8 1863 2 AAQ15363 Pragment
34 563.2 87.7 1251 2 AAQ26981 HCV gene
35 563.2 87.7 3360 2 AAT03677 Hepatitis
36 563.2 87.7 9413 2 AAQ80498 DNA encod
37 563.2 87.7 9413 2 AAQ81559 Hepatitis
38 563.2 87.7 9413 2 AAT03960 Partial H
39 563.2 87.7 9413 6 AAD25517 Hepatitis
40 563.2 87.7 9413 7 AAL53723 Hepatitis
41 563.2 87.7 9413 7 AAD49655 Hepatitis
42 562.2 87.6 2540 2 AAQ29628 Hepatitis
43 560 87.2 1251 2 AAQ25610 HCV in ex
44 560 87.2 1398 2 AAQ22140 Hepatitis
45 559 87.1 742 2 AAQ20926 C10-E15 D


```

Db      421 TATCCGGCCACATAACAGGTACCGTATGCGTTGGGATATGATGATGAACTGTCGCT 480
Qy      481 ACAACGGCCCTGGTGTATCGAGCTGCTCGGATCCCAAGCTGTGCTGGACATGGTG 540
Db      481 ACAACGGCCCTGGTGTATCGAGCTGCTCGGATCCCAAGCTGTGCTGGACATGGTG 540
Qy      541 GCGGGGGCCCAATGGGGAGTCTGGCGGCTCGCTACTATTCCATGCTGGGAACTGG 600
Db      541 GCGGGGGCCCAATGGGGAGTCTGGCGGCTCGCTACTATTCCATGCTGGGAACTGG 600
Qy      601 GCTAAGGTTTGGATTGATGCTACTCTTTGCTCTCTAATAG 642
Db      601 GCTAAGGTTTGGATTGATGCTACTCTTTGCTCTCTAATAG 642

RESULT 3
ADD55634
ID      ADD55634 standard; DNA; 642 BP.
XX
AC      ADD55634;
XX
DT      15-JAN-2004 (first entry)
XX
DE      Hepatitis C virus E1 protein coding sequence #1.
XX
KW      Hepatitis C virus; HCV; vaccine; liver disease; E1 protein; E2 protein;
KW      liver fibrosis; ds; gene.
XX
OS      Hepatitis C virus.
XX
PN      W02003051912-A2.
XX
PD      26-JUN-2003.
XX
PF      18-DEC-2002; 2002WO-EP014480.
XX
PR      18-DEC-2001; 2001US-00020510.
PR      16-OCT-2002; 2002US-0418358P.
XX
PA      (INNO-) INNOGENETICS NV.
XX
PI      Maertens G, Depla B, Bosman F;
XX
DR      WPI; 2003-541632/51.
DR      P-PSDB; ADD55512.
XX
PT      New hepatitis C virus (HCV) vaccine composition, useful for reducing
PT      liver disease, e.g., liver fibrosis in a chronic HCV-infected mammal.
XX
PS      Example 1; SEQ ID NO 3; 271pp; English.
XX
CC      The invention comprises an Hepatitis C virus (HCV) vaccine for reducing
CC      liver disease. The vaccine of the invention comprises an HCV E1 or E2
CC      protein as an antigen. The HCV vaccine is useful for reducing liver
CC      disease (e.g. liver fibrosis) in a chronic HCV-infected mammal. The
CC      present DNA sequence encodes an HCV E1 protein.
XX
SQ      Sequence 642 BP; 109 A; 195 C; 176 G; 162 T; 0 U; 0 Other;

Query Match      100.0%; Score 642; DB 9; Length 642;
Best Local Similarity 100.0%; Pred. No. 2.2e-174;
Matches 642; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 ATGCCCCGTTGCTTTCTTCTATCTCTCTTCTGCTTACTGCTGCTGACCATTTCCA 60
Db      1 ATGCCCCGTTGCTTTCTTCTATCTCTCTTCTGCTTACTGCTGCTGACCATTTCCA 60
Qy      61 GCTTCGCTTATGAGTGCGCAACGTCGCGGATGTACCAATGTCAGCAACGACTGCTCC 120
Db      61 GCTTCGCTTATGAGTGCGCAACGTCGCGGATGTACCAATGTCAGCAACGACTGCTCC 120
Qy      121 AACTCAAGCATTTGTATGAGGACGCGGACATGATCATGCACACCCCGGGTGGTGCCT 180

```

```

Db      121 AACTCAAGCATTTGTATGAGGACGCGGACATGATCATGCACACCCCGGGTGGTGCCT 180
Qy      181 TCGGTTTCGGAGAACAACTCTTTCCGCTGCTGGTAGCGCTCACCCCGGCTCGCAGCT 240
Db      181 TCGGTTTCGGAGAACAACTCTTTCCGCTGCTGGTAGCGCTCACCCCGGCTCGCAGCT 240
Qy      241 AGGAACGCGGCTGCTCCACGACGACATACGACGCGCTGATGCTGCTGGGGG 300
Db      241 AGGAACGCGGCTGCTCCACGACGACATACGACGCGCTGATGCTGCTGGGGG 300
Qy      301 GCTGCTCTCTGTTCCGCTATGTACGTTGGGGATCTCTCGGATCTGCTTCTCGTCTCC 360
Db      301 GCTGCTCTCTGTTCCGCTATGTACGTTGGGGATCTCTCGGATCTGCTTCTCGTCTCC 360
Qy      361 CAGCTGTTTACCATCTCGCTCGCGGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
Db      361 CAGCTGTTTACCATCTCGCTCGCGGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
Qy      421 TATCCGGCCACATTAACAGGTACCGTATGCGTTGGGATATGATGATGAACTGTCGCT 480
Db      421 TATCCGGCCACATTAACAGGTACCGTATGCGTTGGGATATGATGATGAACTGTCGCT 480
Qy      481 ACAACGGCCCTGGTGTATCGAGCTGCTCGGATCCCAAGCTGTGCTGGACATGGTG 540
Db      481 ACAACGGCCCTGGTGTATCGAGCTGCTCGGATCCCAAGCTGTGCTGGACATGGTG 540
Qy      541 GCGGGGGCCCAATGGGGAGTCTGGCGGCTCGCTACTATTCCATGCTGGGAACTGG 600
Db      541 GCGGGGGCCCAATGGGGAGTCTGGCGGCTCGCTACTATTCCATGCTGGGAACTGG 600
Qy      601 GCTAAGGTTTGGATTGATGCTACTCTTTGCTCTCTAATAG 642
Db      601 GCTAAGGTTTGGATTGATGCTACTCTTTGCTCTCTAATAG 642

RESULT 4
AAT12705
ID      AAT12705 standard; DNA; 795 BP.
XX
AC      AAT12705;
XX
DT      23-SEP-1996 (first entry)
XX
DE      HCV E1 construct HCC110A.
XX
KW      HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human;
KW      serotype; reversed phase hybridisation assay; genotype; antigen; sera;
XX
OS      Hepatitis C virus.
XX
PN      W09604385-A2.
XX
PD      15-FEB-1996.
XX
PF      31-JUL-1995; 95WO-EP003031.
XX
PR      29-JUL-1994; 94EP-00870132.
XX
PI      (INNO-) INNOGENETICS NV.
XX
PI      Maertens G, Bosman F, De Martynoff G, Buyse M;
XX
DR      WPI; 1996-129401/13.
XX
PT      Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope proteins
PT      - in presence of dl:sulphide bond cleavage agent, to produce proteins
XX
PS      suitable for direct use in vaccines or diagnostic assays of HCV.
XX
PS      Claim 23; Fig 21; 146pp; English.
XX
CC      AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1
CC      and E2 protein coding constructs. These sequences are included

```

CC in vectors for the production of recombinant E1, E2, and E1/E2 proteins.
CC The recombinant proteins can then be isolated using a method of the
CC invention. In the method, the envelope proteins are purified by carrying
CC out a disulphide bond cleavage, or a reduction step with a disulphide
CC bond cleavage agent, after lysis of recombinant host cells. The
CC constructs containing the purified HCV envelope proteins can be used for
CC vaccinating humans against HCV, for in vitro detection of HCV antibodies
CC in a sample, and in a serotyping assay for detecting one or more
CC serological types of HCV present in a biological sample. The constructs
CC can also be immobilised on a solid substrate and incorporated into a
CC reversed phase hybridisation assay for determining the presence or the
CC genotype of HCV. The new purification method preserves the conformation
CC of the recombinantly expressed E1, E2 and E1/E2, and eliminates
CC contaminating proteins. Antigens isolated using this method are more
CC reactive with human sera than those isolated by known techniques
XX
SQ Sequence 795 BP; 130 A; 240 C; 231 G; 194 T; 0 U; 0 Other;

Query Match 97.9%; Score 628.2; DB 2; Length 795;
Best Local Similarity 98.8%; Pred. No. 2.2e-170;
Matches 633; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 2 TGCCCGGTGCTCTTCTCTATCTCTCTGCTTCTGCTTTACTGCTGTCACCAATCCAG 61
DB 155 TGCCCGGTGCTCTTCTCTATCTCTCTGCTTCTGCTTTGCTGCTGTCACCGTCCAG 214
QY 62 CTTCCCGCTTATGAGTGCACCAAGTGTCCGGGATGTACCAATGTACCAAGCACTGTCCCA 121
DB 215 CTTCCCGCTTATGAGTGCACCAAGTGTCCGGGATGTACCAATGTACCAAGCACTGTCCCA 274
QY 122 ACTCAAGCATTTGTATGAGCGACGACATGATCATGCAACACCCCGGTCGTCCT 181
DB 275 ACTCAAGCATTTGTATGAGCGACGACATGATCATGCAACACCCCGGTCGTCCT 334
QY 182 CGGTTTCGGGAGAACAACTCTTCCCGCTGCTGGGTAGCGCTCACCCCACTGTCAGCTA 241
DB 335 CGGTTTCGGGAGAACAACTCTTCCCGCTGCTGGGTAGCGCTCACCCCACTGTCAGCTA 394
QY 242 GGAACGCGAGCGTCCCAACCAACAAATACGACGCGACGTCGATTTGCTGTTGGGCGG 301
DB 395 GGAACGCGAGCGTCCCAACCAACAAATACGACGCGACGTCGATTTGCTGTTGGGCGG 454
QY 302 CTGCTCTCTGTTCCGCTATGATGAGGAGTCTCTGCGGATCTCTGCGATCTCTCTCTCTCC 361
DB 455 CTGCTCTCTGTTCCGCTATGATGAGGAGTCTCTGCGGATCTCTCTCTCTCTCTCTCTCC 514
QY 362 AGCTGTTTACCATCTGCGCTCGCGGATGAGAGCGTGCAGACTGCAATTTGCTCAATCT 421
DB 515 AGCTGTTTACCATCTGCGCTCGCGGATGAGAGCGTGCAGACTGCAATTTGCTCAATCT 574
QY 422 ATCCCGGCCACATAACAGGTCAACGATATGCTTGGGATATGATGATGAATGTCGCTA 481
DB 575 ATCCCGGCCACATAACAGGTCAACGATATGCTTGGGATATGATGATGAATGTCGCTA 634
QY 482 CAACGCGCTTGGTATGAGAGTCTCTCGGATCCCAACAGCTGTCGTGACATGCTGG 541
DB 635 CAACGCGCTTGGTATGAGAGTCTCTCGGATCCCAACAGCTGTCGTGACATGCTGG 694
QY 542 CGGCGGCCCATTTGGGAGTCTGCGGCGCTCTGCTCTACTATTCCATGTTGGGAACTGGG 601
DB 695 CGGCGGCCCATTTGGGAGTCTGCGGCGCTCTGCTCTACTATTCCATGTTGGGAACTGGG 754
QY 602 CTAAGGTTTGTATGATGATGATCTCTTTGCTCTCTAATAG 642
DB 755 CTAAGGTTTGTATGATGATGATCTCTTTGCTCTCTAATAG 795

RESULT 5
AAL48914
ID AAL48914 standard; DNA; 795 BP.
XX
AC AAL48914;
XX

DT 24-OCT-2002 (first entry)
XX Hepatitis C virus clone HCC110A E1 protein coding sequence.
DE
XX Hepatitis C virus; HCV; E1 protein; E2 protein; infection; gene;
KW Hepatitis C virus; HCV; E1 protein; E2 protein; infection; gene;
KW virucide; immunostimulant; vaccine; ds.
XX Hepatitis C virus.
OS
XX WO200255548-A2.
PN
XX 18-JUL-2002.
PD
XX 11-JAN-2002; 2002WO-BP000219.
XX
PF 11-JAN-2001; 2001US-0260669P.
XX
PR 30-AUG-2001; 2001US-0315768P.
XX
XX (INNO-) INNOGENETICS NV.
PA Maertens G, Bosman F, Buyse M;
XX P-PSDB; AAO18661.
XX
DR WPI: 2002-599657/64.
DR P-PSDB; AAO18661.
XX
PT New therapeutic vaccine compositions comprising at least one purified
PT recombinant hepatitis C virus (HCV) single or specific oligomeric
PT recombinant envelope protein E1 or E2, useful for immunizing humans from
PT HCV infection.
XX
PS Example 2; Page 161-162; 243pp; English.
XX
CC The present invention relates to new therapeutic vaccine compositions for
CC inducing hepatitis C virus (HCV)-specific antibodies, comprising a
CC composition containing at least one purified recombinant HCV single or
CC specific oligomeric recombinant envelope proteins selected from an E1 and
CC an E2 protein, and optionally a pharmaceutical adjuvant. The vaccines are
CC useful for inducing HCV-specific antibodies or for immunising humans
CC against HCV. The recombinant HCV E1 and/or E2 proteins are useful as
CC vaccines or therapeutics, in HCV screening and confirmatory antibody
CC tests, for raising antibodies, in the preparation of medicament, and for
CC in vitro monitoring of HCV disease or prognosing the response to
CC treatment of patients suffering from HCV infection. The present sequence
CC is a coding sequence described in the exemplification of the invention
XX
SQ Sequence 795 BP; 130 A; 240 C; 231 G; 194 T; 0 U; 0 Other;
Query Match 97.9%; Score 628.2; DB 6; Length 795;
Best Local Similarity 98.8%; Pred. No. 2.2e-170;
Matches 633; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 2 TGCCCGGTGCTCTTCTCTATCTCTCTTGGCTTTACTGCTGTCACCAATCCAG 61
DB 155 TGCCCGGTGCTCTTCTCTATCTCTCTTGGCTTTGCTGCTGTCGACCGTTCAG 214
QY 62 CTTCCCGCTTATGAGTGCACCAAGTGTCCGGGATGTACCAATGTACCAAGCACTGTCCCA 121
DB 215 CTTCCCGCTTATGAGTGCACCAAGTGTCCGGGATGTACCAATGTACCAAGCACTGTCCCA 274
QY 122 ACTCAAGCATTTGTATGAGCGACGACATGATCATGCAACACCCCGGTCGTCCT 181
DB 275 ACTCAAGCATTTGTATGAGCGACGACATGATCATGCAACACCCCGGTCGTCCT 334
QY 182 CGGTTTCGGGAGAACAACTCTTCCCGCTGCTGGGTAGCGCTCACCCCACTGTCAGCTA 241
DB 335 CGGTTTCGGGAGAACAACTCTTCCCGCTGCTGGGTAGCGCTCACCCCACTGTCAGCTA 394
QY 242 GGAACGCGAGCGTCCCAACCAACAAATACGACGCGACGTCGATTTGCTGTTGGGCGG 301
DB 395 GGAACGCGAGCGTCCCAACCAACAAATACGACGCGACGTCGATTTGCTGTTGGGCGG 454
QY 302 CTGCTCTCTGTTCCGCTATGATGAGGAGTCTCTGCGGATCTCTCTCTCTCTCTCTCC 361
DB 455 CTGCTCTCTGTTCCGCTATGATGAGGAGTCTCTGCGGATCTCTCTCTCTCTCTCTCTCC 514
QY 362 AGCTGTTTACCATCTGCGCTCGCGGATGAGAGCGTGCAGACTGCAATTTGCTCAATCT 421
DB 515 AGCTGTTTACCATCTGCGCTCGCGGATGAGAGCGTGCAGACTGCAATTTGCTCAATCT 574
QY 422 ATCCCGGCCACATAACAGGTCAACGATATGCTTGGGATATGATGATGAATGTCGCTA 481
DB 575 ATCCCGGCCACATAACAGGTCAACGATATGCTTGGGATATGATGATGAATGTCGCTA 634
QY 482 CAACGCGCTTGGTATGAGAGTCTCTCGGATCCCAACAGCTGTCGTGACATGCTGG 541
DB 635 CAACGCGCTTGGTATGAGAGTCTCTCGGATCCCAACAGCTGTCGTGACATGCTGG 694
QY 542 CGGCGGCCCATTTGGGAGTCTGCGGCGCTCTGCTCTACTATTCCATGTTGGGAACTGGG 601
DB 695 CGGCGGCCCATTTGGGAGTCTGCGGCGCTCTGCTCTACTATTCCATGTTGGGAACTGGG 754
QY 602 CTAAGGTTTGTATGATGATGATCTCTTTGCTCTCTAATAG 642
DB 755 CTAAGGTTTGTATGATGATGATCTCTTTGCTCTCTAATAG 795

Db 515 AGCTGTTTCAACATCTCGCTCGCCGCGCATGAGACGGTGCAGGACTGCAATTGCTCAATCT 574
 Qy 422 ATCCCGGCCACATAACAGGTCAACGATATGGCTTGGGATATGATGATGAATGCTCGCTTA 481
 Db 575 ATCCCGGCCACATAACAGGTCAACGATATGGCTTGGGATATGATGATGAATGCTCGCTTA 634
 Qy 482 CAACGGCCCTGGTGTATCGAGCTGCTCCGGATCCCAAGAGCTGTCTGTGACATGTTGG 541
 Db 635 CAACGGCCCTGGTGTATCGAGCTGCTCCGGATCCCAAGAGCTGTCTGTGACATGTTGG 694
 Qy 542 CGGGGGCCCATTTGGGGAGTCTCGCGGGCTTCGCTACTATTCCATGATGGGGAACTGGG 601
 Db 695 CGGGGGCCCATTTGGGGAGTCTCGCGGGCTTCGCTACTATTCCATGATGGGGAACTGGG 754
 Qy 602 CTAAGGTTTTGATGTGATGCTACTCTTTTC 632
 Db 755 CTAAGGTTTTGATGTGATGCTACTCTTTTC 785

RESULT 9

AAT12973
 ID AAT12973 standard; DNA; 2086 BP.

AC AAT12973;

DT 24-SEP-1996 (first entry)

XX HCV E1 construct HCC165.

DE HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human;
 KW serotype; reversed phase hybridisation assay; genotype; antigen; sera;
 KW ss.

XX Hepatitis C virus.

XX WO9604385-A2.

XX 15-FEB-1996.

XX 31-JUL-1995; 95WO-EP003031.

XX 29-JUL-1994; 94EP-00870132.

XX (INNO-) INNOGENETICS NV.

XX Maertens G, Boeman F, De Martynoff G, Buyse M;

XX WPI; 1996-129401/13.

XX Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope proteins
 PT - in presence of di-sulphide bond cleavage agent, to produce proteins
 PT suitable for direct use in vaccines or diagnostic assays of HCV.

XX Claim 23; Fig 21; 146pp; English.

XX AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1
 CC and E2 protein coding sequence constructs. These sequences are included
 CC in vectors for the production of recombinant E1, E2, and E1/E2 proteins.
 CC The recombinant proteins can then be isolated using a method of the
 CC invention. In the method, the envelope proteins are purified by carrying
 CC out a disulphide bond cleavage, or a reduction step with a disulphide
 CC bond cleavage agent, after lysis of recombinant host cells. The
 CC constructs containing the purified HCV envelope proteins can be used for
 CC vaccinating humans against HCV, for in vitro detection of HCV antibodies
 CC in a sample, and in a serotyping assay for detecting one or more
 CC serological types of HCV present in a biological sample. The constructs
 CC can also be immobilised on a solid substrate and incorporated into a
 CC reversed phase hybridisation assay for determining the presence or the
 CC genotype of HCV. The new purification method preserves the conformation
 CC of the recombinantly expressed E1, E2 and E1/E2, and eliminates
 CC contaminating proteins. Antigens isolated using this method are more
 CC reactive with human sera than those isolated by known techniques

XX

SQ Sequence 2086 BP; 366 A; 635 C; 601 G; 484 T; 0 U; 0 Other;
 Query Match 96.5%; Score 619.8; DB 2; Length 2086;
 Best Local Similarity 98.9%; Pred. No. 8.3e-168;
 Matches 624; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 2 TGCCCGGTGCTCTTTCTCTATCTTCTCTTGGCTTTTACTGTCCTGTCTGACATTTCCAG 61
 Db 155 TGCCCGGTGCTCTTTCTCTATCTTCTCTTGGCTTTTACTGTCCTGTCTGACATTTCCAG 214
 Qy 62 CTTCGGCTTTATGAGTGGCAACGTCGCGGATGTACCATGTACAGCAACGATCTCTCCA 121
 Db 215 CTTCGGCTTTATGAGTGGCAACGTCGCGGATGTACCATGTACAGCAACGATCTCTCCA 274
 Qy 122 ACTCAAGCATTTGTATGAGCGAGCGGACATGATCATGACACCCCGGCTGCGTCCCT 181
 Db 275 ACTCAAGCATTTGTATGAGCGAGCGGACATGATCATGACACCCCGGCTGCGTCCCT 334
 Qy 182 GCGTTGGGAGAACTCTTCTCTTCCGCTCTGCTGGGTAGCGCTACCCGCCAGCTCGAGCTA 241
 Db 335 GCGTTGGGAGAACTCTTCTCTTCCGCTCTGCTGGGTAGCGCTACCCGCCAGCTCGAGCTA 394
 Qy 242 GGAAGCCAGAGTCCCAACACGCAATAACGACGACGTCGATTTGCTGTTGGGGCG 301
 Db 395 GGAAGCCAGAGTCCCAACACGCAATAACGACGACGTCGATTTGCTGTTGGGGCG 454
 Qy 302 CTGCTCTCTGTTCCGCTATGTACATGGGGGATCTCTGCGGATCTCTTCTCTGCTCC 361
 Db 455 CTGCTCTCTGTTCCGCTATGTACGTTGGGGACCTCTGCGGATCTCTTCTCTGCTCC 514
 Qy 362 AGCTGTTCCATCTCTGCTCGCGCATGAGACGCTGAGCACTGCAATTCGCTCAATCT 421
 Db 515 AGCTGTTCCATCTCTGCTCGCGCATGAGACGCTGAGCACTGCAATTCGCTCAATCT 574
 Qy 422 ATCCCGGCCACATAACAGGTCAACGTCATGGCTTTGGGATATGATGATGATGATGATGAT 481
 Db 575 ATCCCGGCCACATAACAGGTCAACGTCATGGCTTTGGGATATGATGATGATGATGATGAT 634
 Qy 482 CAACGGCCCTGGTGTATCGAGCTGCTCCGGATCCCAAGCTGTCGTGACATGCTGG 541
 Db 635 CAACGGCCCTGGTGTATCGAGCTGCTCCGGATCCCAAGCTGTCGTGACATGCTGG 694
 Qy 542 CGGGGGCCCATTTGGGGAGTCTCTGCGGGCTTCGCTACTATTCCATGATGGGGAACTGGG 601
 Db 695 CGGGGGCCCATTTGGGGAGTCTCTGCGGGCTTCGCTACTATTCCATGATGGGGAACTGGG 754
 Qy 602 CTAAGGTTTTGATGTGATGCTACTCTTTTC 632
 Db 755 CTAAGGTTTTGATGTGATGCTACTCTTTTC 785

RESULT 10

AAT12974
 ID AAT12974 standard; DNA; 2433 BP.

XX AC AAT12974;

XX DT 25-SEP-1996 (first entry)

XX DE HCV E1 construct HCC166.

XX HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human;
 KW serotype; reversed phase hybridisation assay; genotype; antigen; sera;
 KW ss.

XX Hepatitis C virus.

XX WO9604385-A2.

XX 15-FEB-1996.

XX 31-JUL-1995; 95WO-BP003031.

XX

[illegible]

RESULT 12

RESUB 12
AAI.48940

AAI48940
ID AAI48940 standard: DNA: 2434 BP

XX
IT
AA40340

AC AAT.48940:

[illegible]

DT 24-OCT-2002 (first entry)

XX

DE Hepatitis C virus E2 protein related coding sequence SEQ ID NO:

[illegible]

Hepatitis C virus: HCV. E1 protein. E2 protein. infection. gene.

KW hepatitis C virus; HCV; E1 protein; E2 protein; gene;
KW virocyde; immunostimulant; vaccine; ds

virucide; immunostimulant; vaccine; ds.
 xw
 yy

20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525
 526
 527
 528
 529
 530
 531
 532
 533
 534
 535
 536
 537
 538
 539
 540
 541

Hepatitis C virus.

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

PN WO200255548-A2.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

PD 18-JUL-2002.

XX

PF 11-JAN-2002; 2002WO-EP000219.



PR 11-JAN-2001; 2001US-0260669P.

PR 30-AUG-2001; 2001US-0315768P.

XX

PA (INNO-) INNOGENETICS NV.

[illegible]

PI Maertens G. Borsman F. Bilyaev M:

[illegible]

useful for inducing HCV-specific antibodies or for immunising humans against HCV. The recombinant HCV E1 and/or E2 proteins are useful as vaccines or therapeutics, in HCV screening and confirmatory antibody tests, for raising antibodies, in the preparation of medicament, and for in vitro monitoring of HCV disease or prognosing the response to treatment of patients suffering from HCV infection. The present sequence is a coding sequence described in the exemplification of the invention

Query Match	94.9%	Score 608.8	DB 6	Length 2434
Best Local Similarity	98.7%	Pred. No. 1.3e-164		
Matches 624	Conservative 0	Mismatches 7	Indels 1	Gaps 1
Qy	2	TGCCCGGTTCCTCTTCTCTATCTCTCTCTTGGCTTTTACTGTCCTG-TCTGACCAATTCCA	60	
Db	506	TGCCCGGTTCCTCTTCTCTATCTCTCTCTTGGCTTTGCTGCTGTTCTGACCGTTCCA	565	
Qy	61	GCTTCGCTTATGAGGTGCGCAACGTGTCCGGGATGACCATGTCTACGAAAGCATGCTCC	120	
Db	566	GCTTCGCTTATGAGTGGCAAGGTGTCCGGGATGACCATGTCTACGAAAGCATGCTCC	625	
Qy	121	AACCTAAGCAATTGTGTATGAGGAGCGGACATGATCATGACACACCCCGGGTGCCTGCC	180	
Db	626	AACCTAAGCAATTGTGTATGAGGAGCGGACATGATCATGACACACCCCGGGTGCCTGCC	685	
Qy	181	TGCGTTGGGAGAACAACTCTTCCGCTCTGGGTAGCGCTCACCCCAAGCTCGCAGCT	240	
Db	686	TGCGTTGGGAGAACAACTCTTCCGCTCTGGGTAGCGCTCACCCCAAGCTCGCAGCT	745	
Qy	241	AGGAACCGCAGCGTCCCAACACAGCAATACAGACGCCAGTTCGATTTGCTCGTTGGGGCG	300	
Db	746	AGGAACCGCAGCGTCCCAACACAGCAATACAGACGCCAGTTCGATTTGCTCGTTGGGGCG	805	
Qy	301	GCTGCTCTCTGTTCCGCTATGTACGTGGGGGATCTCTGCGGATCTGCTTCTCTCGTCTCC	360	
Db	806	GCTGCTCTCTGTTCCGCTATGTACGTGGGGGATCTCTGCGGATCTGCTTCTCTCGTCTCC	865	
Qy	361	CAGCTGTTTCCACATCTCGGCTCGCCGGCATGAGACGGTGCAGGACATGCTCAATTC	420	
Db	866	CAGCTGTTTCCACATCTCGGCTCGCCGGCATGAGACGGTGCAGGACATGCTCAATTC	925	
Qy	421	TATCCCGGCCACATAACAGCTCACCGTATGGCTTGGGATATGATGATGAACCTGGTCGCCT	480	
Db	926	TATCCCGGCCACATAACGGGTCAACCGTATGGCTTGGGATATGATGATGAACCTGGTCGCCT	985	
Qy	481	ACAAAGCCGCTGGTGGTATTCGACGCTGCTCCGATCCCAAGCTGTCTGTCGACATGGTG	540	
Db	986	ACAAAGCCGCTGGTGGTATTCGACGCTGCTCCGATCCCAAGCTGTCTGTCGACATGGTG	1045	
Qy	541	GGCGGGGCCCATTTGGGAGTCTTGCGGGCCCTCGCCTACTATTTCATGGTGGGGAACCTGG	600	
Db	1046	GGCGGGGCCCATTTGGGAGTCTTGCGGGCCCTCGCCTACTATTTCATGGTGGGGAACCTGG	1105	
Qy	601	GCTAAGGTTTGAATTGTGATGCTACTCTTTGCG	632	
Db	1106	GCTAAGGTTTGGTTGTGATGCTACTCTTTGCG	1137	

PRISM.T 13

RESULTS 13
AA064069

AAQ64068
IN AAQ64068 standard: CDA: 3461 BP

AAQ64068

5000

AC
AAQ64068;

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

DT 14-FEB-19

DE Non-A, no

XX

KW Non-A, no

KW NS1; NS2;


```

Db      1327 CAACAGCCCTAGTGTGTCTGCGAGTTGCTCCGATCCCAAGCTGTGTCGACATGTTGG 1386
Qy      542  CGGGGGCCCATTTGGGGAGTCTCTGGGGCCCTCGCCTACTATTCCATGGTGGGNACTGGG 601
Db      1387 CGGGGGCCCACTGGGGAGTCTCTGGGGCCCTTGCCTACTATTCCATGGTAGGGAACTGGG 1446
Qy      602  CTAAGGTTTGTGATTGTGANGCTACTCTTTGC 632
Db      1447 CTAAGGTCCTGATTGTGGCGTACTCTTTGC 1477

```

Search completed: June 16, 2004, 09:36:51
 Job time : 244.553 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2004, 09:12:53 ; Search time 1725.91 Seconds
(without alignments)
11108.065 Million cell updates/sec

Title: US-09-899-303a-3
Perfect score: 642
Sequence: 1 ATGCCGGTGTCTTCTC.....TACTCTTGTCTCTAATAG 642

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- EST:*
- 1: em_estba:*
 - 2: em_esthum:*
 - 3: em_estin:*
 - 4: em_estmu:*
 - 5: em_estov:*
 - 6: em_estpl:*
 - 7: em_estro:*
 - 8: em_hic:*
 - 9: gb_est1:*
 - 10: gb_est2:*
 - 11: gb_hic:*
 - 12: gb_est3:*
 - 13: gb_est4:*
 - 14: gb_est5:*
 - 15: em_estom:*
 - 16: em_estom:*
 - 17: em_gss_hum:*
 - 18: em_gss_inv:*
 - 19: em_gss_pln:*
 - 20: em_gss_vrt:*
 - 21: em_gss_fun:*
 - 22: em_gss_mam:*
 - 23: em_gss_mus:*
 - 24: em_gss_pro:*
 - 25: em_gss_rod:*
 - 26: em_gss_phg:*
 - 27: em_gss_vrl:*
 - 28: gb_gss1:*
 - 29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
C 1	117.2	18.3	488	9	AV755731
C 2	92.6	14.4	492	9	AV758366
C 3	41.6	6.5	534	14	CF846043
C 4	40.6	6.3	925	29	CNS0091P

5	40.4	6.3	590	14	CA659369
6	40.4	6.3	645	29	CNS01213
7	40	6.2	525	10	BE337089
8	40	6.2	671	12	BI723733
9	40	6.2	740	12	BJ536071
10	40	6.2	1201	13	BX356664
11	39.6	6.2	423	9	AA459034
12	39	6.1	621	14	CA816001
13	39	6.1	624	12	BI723734
14	38.6	6.0	399	9	AV638521
15	38.6	6.0	434	9	AV637507
16	38.6	6.0	440	9	AV637983
17	38.6	6.0	450	9	AV637259
18	38.6	6.0	451	9	AV637328
19	38.6	6.0	451	9	AV637643
20	38.6	6.0	453	9	AV634724
21	38.6	6.0	454	9	AV637050
22	38.6	6.0	456	9	AV635382
23	38.6	6.0	473	9	AV632765
24	38.6	6.0	481	9	AV635503
25	38.6	6.0	485	9	AV632811
26	38.6	6.0	506	9	AV392445
27	38.6	6.0	508	9	AV634095
28	38.6	6.0	526	9	AV641895
29	38.6	6.0	533	9	AV638125
30	38.6	6.0	537	9	AV632335
31	38.6	6.0	588	9	AV387329
32	38.6	6.0	693	13	CA160850
33	38.2	6.0	431	9	AV639153
34	38.2	6.0	501	9	AV638474
35	38.2	6.0	1148	14	CK208676
C 36	38	5.9	742	13	BQ752673
C 37	38	5.9	1195	28	E10902
C 38	37.8	5.9	435	13	C72860
C 39	37.8	5.9	634	10	BF396591
40	37.4	5.8	497	9	AV633658
41	37.4	5.8	655	29	CG274746
42	37.4	5.8	782	29	CC605275
43	37.4	5.8	856	28	BZ578381
44	37.4	5.8	860	29	CG286130
45	37.4	5.8	872	28	BZ555011

ALIGNMENTS

RESULT 1	AV755731/c	AV755731	488 bp	mRNA	linear	EST 19-OCT-2000
LOCUS	AV755731	BM Homo sapiens	CDNA clone	BMFAKB03	5', mRNA sequence.	
DEFINITION	AV755731	BM Homo sapiens	CDNA clone	BMFAKB03	5', mRNA sequence.	
ACCESSION	AV755731	EST.				
VERSION	AV755731.1	GI:10913579				
KEYWORDS	EST.					
SOURCE	Homo sapiens (human)					
ORGANISM	Homo sapiens					
REFERENCE	Sukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
AUTHORS	Gu, J., Zhao, M., Huang, Q., Xu, X., Li, Y., Peng, Y., Song, H., Xiao, H., Gu, Y., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z., Zeng, B., Xu, S., Gu, W., Tu, Y., Jie, J., Fu, G., Ren, S., Zhong, M., Lu, G., Yang, Y., Gao, G., Wang, Z., Zhang, Q., Chen, S., Han, Z., and Chen, Z.					
TITLE	Homo sapiens CDNA BM clones					
JOURNAL	Unpublished (2000)					
COMMENT	Contact: Zequang Han Chinese National Human Genome Center at Shanghai 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China Tel: 86-21-50801919 (ex. 45) Fax: 86-21-50801922 Email: hanzg@chgc.sh.cn This clone is available at CHGC in Shanghai.					


```
ORIGIN
Query Match      6.5%; Score 41.6; DB 14; Length 534;
Best Local Similarity 47.1%; Pred. No. 20;
Matches 128; Conservative 0; Mismatches 144; Indels 0; Gaps 0;

QY 70 TATGAGTGGCGCAACGTGTCCGGATGTACATCTCAGAACGACTGCTCAACTCAAGC 129
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 200 TACGCGGTGGCGGAGATTACGGTATCCGATGCGCTTCGCCGCTTCTACAACTGGACC 259
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 130 ATTGTGTATGAGGCGAGCATATCATGACACACCCCGGGTGGCTGCTGGTTCGG 189
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 260 TCATGACACAGAGAGAGGCCCGCATCATGCTGACCCCAAGACGGTGGCCAAATTCAC 319
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 190 GAGAACAACTCTTCCCGCTGCTGGGTAGCGCTCACCCCGCTCGAGCTAGGAAGCC 249
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 320 CACTACGGCGGCACCATCTCTGGCTCGAACCGTGGTGGCTTCGACGTGGACAGATTATC 379
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 250 AGCGTCCCAACACGACAAATACGACGCAACGTGATTTGCTGTGGGGGGCTGCTCTC 309
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 380 AACTTCTCAGACGACGACGCGCTCTCGAGGTATACGTGATCGCGGTGACGGCAACCCAC 439
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 310 TGTTCGGCTATGTACGTTGGGGGATCTCTGGG 341
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 440 CGTGGCGCCAAAGATCTCGAGAGGTGCGG 471
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
RESULT 4
CNS0091P
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence TET3 end of BAC #
BACR19D16 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION
AL053013
VERSION
AL053013.1 GI:4934461
KEYWORDS
GSS.
SOURCE
Drosophila melanogaster (fruit fly)
ORGANISM
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Phytophaga; Drosophilidae; Drosophila.
1 (bases 1 to 925)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
EP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammeter in Peter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
F1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
Location/Qualifiers
1..925
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
```

```
ORIGIN
Query Match      6.3%; Score 40.6; DB 29; Length 925;
Best Local Similarity 15.8%; Pred. No. 45;
Matches 69; Conservative 165; Mismatches 198; Indels 4; Gaps 1;

QY 122 ACTCAAGCATTTGTATGAGGCGAGCATGATCATGCACACCCCGGTCGTCGCTT 181
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 489 AATANANNNTTATTTATTAANNNNANANANANNNNNAGCSMSGCKKCGSTTBS 548
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 182 CGGTTCGGAGAGCAAACTCTTCCCGCTGCTGGGTAGCGCTCACCCCGGTCGTCG 241
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 549 TTTTTSAGSGYGRGSGSGSCSCSSCSCSSSCSCCSCCSCSSYCCSSBSBS 608
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 242 GGAAGCCAGGTCCTCCACACGACATACAGCCAGCTGATGCTGTTGGGGGG 301
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 609 KCSSTBSBSCSCSSKSVCGTSCSS---SSSCSSSSSTSSSTSSSTSSKSSSS 664
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 302 CTGCTCTCTGTTCCGCTATGTACGTGGGGGATCTCTGCGGATCTCTTCTCTGCTCC 361
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 665 SSSSTTTSKTSASAGSGSWAGSGSGSTGTSSTSSSSSTSSSTSSSSSSSSBS 724
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 362 AGCTGTTTCAACATCTCGCTCGCGCATGAGAGCGGTGACGACTGCAATGCTCAATCT 421
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 725 BSSSGSSSSSTSSBBSCTSTSSSSSSSYSSSTSCCTCCCSYSSSTSSSTSSW 784
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 422 ATCCCGGCGCAATACAGGTACCGTATGCTTGGGATATGATGATGATGATGATG 481
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 785 SGSSSSSVGTSSSDSTSTCCSCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 844
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 482 CAACGCGCTGCTGCTATCGCAGCTCTCCGATCCACAGCTCTCTCTGTCGATGCTG 541
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 845 CGCGSSSTGMBGTSSACSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 904
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 542 CGGGGGCCCATTTGGG 557
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 905 KSSSSGSGSVSSGSGSGS 920
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 5
CA659369
LOCUS
DEFINITION
wlm1.pk0005.c9 wlm1 Triticum aestivum cDNA clone wlm1.pk0005.c9 5'
end, mRNA sequence.
ACCESSION
CA659369
VERSION
CA659369.1 GI:25237894
KEYWORDS
EST.
SOURCE
Triticum aestivum (bread wheat)
ORGANISM
Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooideae; Triticeae; Triticum.
1 (bases 1 to 590)
Tingey, S.V., Powell, W., Wolters, P., Dolan, M., Hainey, C., Yuan, Z.,
Miao, G., Caraher, N. and Hanafey, M.K.
DUPONT Wheat cDNA Sequence
Unpublished (2002)
Contact: Scott V. Tingey
Crop Genetics
E. I. DuPont de Nemours and Company
1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA
Tel: 302-631-2602
Fax: 302-631-2607
Email: Scott.V.Tingey@USA.dupont.com
Seq primer: M13.
Location/Qualifiers
1..590
/organism="Triticum aestivum"
/mol_type="mRNA"

Search completed: June 16, 2004, 13:27:08
Job time : 1729.91 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2004, 09:14:33 ; Search time 46.2775 Seconds
(without alignments)
7698.741 Million cell updates/sec

Title: US-09-899-303A-3

Perfect score: 642

Sequence: 1 ATCCCGGTGCTCTTCTC.....TACTCTTGTCTCTCTAATAG 642

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*

1: /cgn2_6/ptodata/2/ina/5A COMB.seq.*

2: /cgn2_6/ptodata/2/ina/5B COMB.seq.*

3: /cgn2_6/ptodata/2/ina/6A COMB.seq.*

4: /cgn2_6/ptodata/2/ina/6B COMB.seq.*

5: /cgn2_6/ptodata/2/ina/ACTUS COMB.seq.*

6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	642	100.0	642	3	US-08-612-973-3
2	642	100.0	642	3	US-08-927-597-3
3	628.2	97.9	795	3	US-08-612-973-5
4	628.2	97.9	795	3	US-08-927-597-5
5	619.8	96.5	2082	3	US-08-612-973-47
6	619.8	96.5	2082	3	US-08-927-597-47
7	619.8	96.5	2433	3	US-08-612-973-49
8	619.8	96.5	2433	3	US-08-927-597-49
9	565.4	88.1	11076	4	US-09-539-601-19
10	565.4	88.1	11076	4	US-09-539-601-19
11	565.4	88.1	11076	4	US-09-539-601-25
12	565.4	88.1	11076	4	US-09-539-601-31
13	563.8	87.8	1539	2	US-08-470-426B-17
14	563.8	87.8	1863	2	US-08-470-426B-14
15	559	87.1	742	1	US-08-081-072-18
16	559	87.1	742	1	US-08-449-093A-18
17	559	87.1	932	1	US-08-081-072-15
18	559	87.1	932	1	US-08-449-093A-15
19	555.8	86.6	2116	3	US-08-191-160-21
20	555.8	86.6	9595	3	US-09-014-416-4
21	555.8	86.6	9599	3	US-09-014-416-6
22	544.6	84.8	9472	4	US-08-150-204B-96
23	540.8	84.2	1167	1	US-08-324-977-9
24	540.8	84.2	1167	2	US-08-384-616-9
25	540.8	84.2	1167	2	US-08-904-686A-9
26	540.8	84.2	1167	3	US-09-315-850-9
27	540.8	84.2	1499	1	US-08-324-977-3

28	540.8	84.2	1499	2	US-08-384-616-3	Sequence 3, Appli
29	540.8	84.2	1499	2	US-08-904-686A-3	Sequence 3, Appli
30	540.8	84.2	1499	2	US-09-315-850-3	Sequence 3, Appli
31	540.8	84.2	6039	1	US-08-324-977-11	Sequence 11, Appl
32	540.8	84.2	6039	2	US-08-384-616-11	Sequence 11, Appl
33	540.8	84.2	6039	2	US-08-904-686A-11	Sequence 11, Appl
34	540.8	84.2	6039	3	US-09-315-850-11	Sequence 11, Appl
35	540.8	84.2	9030	1	US-08-324-977-13	Sequence 13, Appl
36	540.8	84.2	9030	2	US-08-384-616-13	Sequence 13, Appl
37	540.8	84.2	9030	2	US-08-904-686A-13	Sequence 13, Appl
38	540.8	84.2	9030	3	US-09-315-850-13	Sequence 13, Appl
39	540.8	84.2	9416	1	US-08-324-977-1	Sequence 1, Appli
40	540.8	84.2	9416	2	US-08-384-616-1	Sequence 1, Appli
41	540.8	84.2	9416	2	US-08-904-686A-1	Sequence 1, Appli
42	540.8	84.2	9416	3	US-09-315-850-1	Sequence 1, Appli
43	540.8	84.2	9416	4	US-08-823-895A-27	Sequence 27, Appl
44	511.8	79.7	576	1	US-08-086-428B-21	Sequence 21, Appl
45	511.8	79.7	576	2	US-08-468-570-21	Sequence 21, Appl

ALIGNMENTS

RESULT 1
US-08-612-973-3
; Sequence 3, Application US/08612973
; Patent No. 6150334
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, PONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612.973
; FILING DATE: 11-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 642 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..639
; FEATURE:
; NAME/KEY: mat peptide
; LOCATION: 1..636

[illegible]

RESULT 9

```

US-09-539-601-1
; Sequence 1, Application US/09539601C
; Patent NO. 6630343
; GENERAL INFORMATION:
; APPLICANT: Bartenschlager, Ralf FW
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System
; FILE REFERENCE: all sequences
; CURRENT APPLICATION NUMBER: US/09/539,601C
; CURRENT FILING DATE: 2001-08-30
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY
; EARLIER FILING DATE: 1999-04-03
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 11076
; TYPE: DNA
; ORGANISM: Hepatitis C virus
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (1)..(341)
; OTHER INFORMATION: construct I389/Core-3'/wt
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (342)..(1193)
; OTHER INFORMATION: hepatitis C virus core - neomycin
; OTHER INFORMATION: phosphotransferase fusion protein
; FEATURE:
; NAME/KEY: RBS
; LOCATION: (1202)..(1812)
; OTHER INFORMATION: internal ribosome entry site from
; OTHER INFORMATION: encephalomyocarditis virus
; FEATURE:

```

```

; NAME/KEY: CDS
; LOCATION: (1813)..(10845)
; OTHER INFORMATION: hepatitis C virus polypeptide from core to
; OTHER INFORMATION: nonstructural protein NS5B; parental sequence
; OTHER INFORMATION: without cell culture-adaptive mutations
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: (10846)..(11076)
US-09-539-601-1

```

[illegible]

RESIST 10

```

: RESOUR 10
: US-09-539-601-19
: Sequence 19, Application US/09539601C
: Patent No. 6630343
: GENERAL INFORMATION:
: APPLICANT: Bartenschlager, Ralf FW
: TITLE OF INVENTION: Hepatitis C Virus Cell Culture System
: FILE REFERENCE: all sequences
: CURRENT APPLICATION NUMBER: US/09/539,601C
: CURRENT FILING DATE: 2001-08-30
: EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY
: EARLIER FILING DATE: 1999-04-03
: NUMBER OF SEQ ID NOS: 51
: SOFTWARE: PatentIn Ver. 2.1

```


SUMMARIES

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2004, 08:32:08 ; Search time 2425.41 Seconds
(without alignments)
14206.949 Million cell updates/sec

Title: US-09-899-303A-5
Perfect score: 795
Sequence: 1 ATGTTGGTAAGTCAATCGA.....TACTCTTGTCTCCTAATAG 795

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pin.*
- 35: em_htg_rod.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	795	100.0	795	6 A48667	A48667 Sequence 5
2	795	100.0	795	6 ARI57325	ARI57325 Sequence
3	795	100.0	795	6 AX452754	AX452754 Sequence
4	795	100.0	795	6 AX685006	AX685006 Sequence
5	795	100.0	795	6 AX794848	AX794848 Sequence
6	778.8	98.0	2082	6 A48709	A48709 Sequence 47
7	778.8	98.0	2082	6 ARI57350	ARI57350 Sequence
8	778.8	98.0	2082	6 AX452796	AX452796 Sequence
9	778.8	98.0	2082	6 AX685048	AX685048 Sequence
10	778.8	98.0	2082	6 AX794890	AX794890 Sequence
11	778.8	98.0	2433	6 A48711	A48711 Sequence 49
12	778.8	98.0	2433	6 ARI57351	ARI57351 Sequence
13	778.8	98.0	2433	6 AX452798	AX452798 Sequence
14	778.8	98.0	2433	6 AX685050	AX685050 Sequence
15	778.8	98.0	2433	6 AX794892	AX794892 Sequence
16	711.6	89.5	9379	14 AF207766	AF207766 Hepatitis
17	711.6	89.5	9379	14 HCVPOLYP	AJ7000009 Hepatitis
18	710	89.3	9386	14 AF165056	AF165056 Hepatitis
19	710	89.3	9386	14 AF165056	AF165056 Hepatitis
20	708.4	89.1	9386	14 AF165055	AF165055 Hepatitis
21	708.4	89.1	9410	14 HPCIR2	D50481 Hepatitis C
22	706.8	88.9	3296	14 AB008441	AB008441 Hepatitis
23	706.8	88.9	3296	14 AB008443	AB008443 Hepatitis
24	706.8	88.9	9379	14 AF165052	AF165052 Hepatitis
25	706.8	88.9	9379	14 AF207761	AF207761 Hepatitis
26	705.2	88.7	3296	14 AB008442	AB008442 Hepatitis
27	705.2	88.7	3296	14 AB008446	AB008446 Hepatitis
28	705.2	88.7	9379	14 AF165051	AF165051 Hepatitis
29	705.2	88.7	9598	14 AB049101	AB049101 Hepatitis
30	703.6	88.5	3296	14 AB008447	AB008447 Hepatitis
31	703.6	88.5	9547	14 AB049091	AB049091 Hepatitis
32	703.6	88.5	9605	6 AX739971	AX739971 Sequence
33	703.6	88.5	9605	14 HCV238799	AJ238799 Hepatitis
34	703.6	88.5	11076	6 AR406040	AR406040 Sequence
35	703.6	88.5	11076	6 AR406046	AR406046 Sequence
36	703.6	88.5	11076	6 AR406048	AR406048 Sequence
37	703.6	88.5	11076	6 AR406050	AR406050 Sequence
38	703.6	88.5	11076	6 AX036252	AX036252 Sequence
39	703.6	88.5	11076	6 AX036258	AX036258 Sequence
40	703.6	88.5	11076	6 AX036260	AX036260 Sequence
41	703.6	88.5	11076	6 AX036262	AX036262 Sequence
42	702	88.3	1880	14 HPC5TRJ4	D00832 Hepatitis C
43	702	88.3	2540	6 E04260	E04260 cDNA encodi
44	702	88.3	2540	6 E04805	E04805 cDNA to 5'-
45	702	88.3	2540	6 E07391	E07391 cDNA encodi

ALIGNMENTS

RESULT 1

A48667	A48667	795 bp	DNA	linear	PAT 07-MAR-1997
LOCUS	Sequence 5 from Patent WO9604385.				
DEFINITION	A48667				
ACCESSION	A48667.1	GI:2302380			
VERSION	unidentified				
KEYWORDS	unidentified				
SOURCE	unclassified.				
ORGANISM	1 (bases 1 to 795)				
REFERENCE	Maertens,G., Bosman,F., De,M.G. and Buyse,M.				
AUTHORS	PURIFIED HEPATITIS C VIRUS ENVELOPE PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE				
TITLE	Patent: WO 9604385-A 5 15-FEB-1996;				
JOURNAL					

INNOGENETICS NV (BE)
Other publication CA 2172273 960215
Other publication AU 3382495 960304.
Location/Qualifiers
1. .795
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"
1. .792
/note="unnamed protein product"
/codon_start=1
/protein_id="CAA03129.1"
/db_xref="GI:2302381"
/translation="MLGKVIDTLTCGFADLVGIPLVGAPLGAARALAHGVRVLEDD
VNYATGNLPGCSFSLPALLSCLTPASAVEVRNVSGMVHTVNDSCNSIVVRAADM
IMHTPCVPCVBNSSRCVVALTPTLAARNAVPTTIRRHVDLLVGAARAFCSAMV
GDLCSVFLVLSQFTISPRHETVQDCNSIYPGHTIGRMAWMDMMWSPITALVVS
QLLRIPQAVDVMVAGAHGVLGAGLAYTSMVGNWAKVLIYMLLPAP"
1. .789
/product="unnamed"
mat_peptide
ORIGIN
Query Match 100.0%; Score 795; DB 6; Length 795;
Best Local Similarity 100.0%; Pred. No. 3e-167;
Matches 795; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGTGGTGAAGTCAATGATACCTTACATCGGCTTCGCCGACCTCGTGGGGTACATT 60
DB 1 ATGTGGTGAAGTCAATGATACCTTACATCGGCTTCGCCGACCTCGTGGGGTACATT 60
QY 61 CCGCTCGTGGGGCCCCCTAGGGGGCGCTGCGAGGGCCCTGCGCATGCGGTCGCGGTT 120
DB 61 CCGCTCGTGGGGCCCCCTAGGGGGCGCTGCGAGGGCCCTGCGCATGCGGTCGCGGTT 120
QY 121 CTGGAGGACGGGTGAATATGCAACAGGGAAATTTGCCCCGGTTCCTTCTCTATCTTC 180
DB 121 CTGGAGGACGGGTGAATATGCAACAGGGAAATTTGCCCCGGTTCCTTCTCTATCTTC 180
QY 181 CTCTGGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
DB 181 CTCTGGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
QY 241 TCGGGATGTACCATGTCAAGACGACTGCTCAACTCAAGCATTTGTTATGAGGAGCG 300
DB 241 TCGGGATGTACCATGTCAAGACGACTGCTCAACTCAAGCATTTGTTATGAGGAGCG 300
QY 301 GACATGATCATGACACACCCCGGGTGGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
DB 301 GACATGATCATGACACACCCCGGGTGGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
QY 361 TGCTGGGTAGCGCTACACCCCGGGTGGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
DB 361 TGCTGGGTAGCGCTACACCCCGGGTGGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
QY 421 ATACGACGCGCAGCTGCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
DB 421 ATACGACGCGCAGCTGCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
QY 481 GGGGACCTCTGCGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
DB 481 GGGGACCTCTGCGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
QY 541 CATGACGCGTGCAGGACTGCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
DB 541 CATGACGCGTGCAGGACTGCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
QY 601 ATGGCTTGGGATATGATGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
DB 601 ATGGCTTGGGATATGATGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
QY 661 CTCGGATCCCAAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
DB 661 CTCGGATCCCAAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720

DB 661 CTCGGATCCCAAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
QY 721 GGTCTCGCTACTATTCCCATGCTGGGAACTGGGCTAAAGCTTTTGTATGATGCTACTC 780
DB 721 GGTCTCGCTACTATTCCCATGCTGGGAACTGGGCTAAAGCTTTTGTATGATGCTACTC 780
QY 781 TTTGCTCCCTAATAG 795
DB 781 TTTGCTCCCTAATAG 795
RESULT 2
AR157325
LOCUS AR157325 795 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 5 from patent US 6245503
ACCESSION AR157325
VERSION AR157325.1 GI:16218258
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 795)
AUTHORS Maertens, G., Bosman, F., De Martynoff, G. and Buysse, M.-A.
TITLE Purified hepatitis C virus envelope proteins for diagnostic and therapeutic use
JOURNAL Patent: US 6245503-A 5 12-JUN-2001;
FEATURES
source 1. .795
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Query Match 100.0%; Score 795; DB 6; Length 795;
Best Local Similarity 100.0%; Pred. No. 3e-167;
Matches 795; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGTGGTGAAGTCAATGATACCTTACATCGGCTTCGCCGACCTCGTGGGGTACATT 60
DB 1 ATGTGGTGAAGTCAATGATACCTTACATCGGCTTCGCCGACCTCGTGGGGTACATT 60
QY 61 CCGCTCGTGGGGCCCCCTAGGGGGCGCTGCGAGGGCCCTGCGCATGCGGTCGCGGTT 120
DB 61 CCGCTCGTGGGGCCCCCTAGGGGGCGCTGCGAGGGCCCTGCGCATGCGGTCGCGGTT 120
QY 121 CTGGAGGACGGGTGAATATGCAACAGGGAAATTTGCCCCGGTTCCTTCTCTATCTTC 180
DB 121 CTGGAGGACGGGTGAATATGCAACAGGGAAATTTGCCCCGGTTCCTTCTCTATCTTC 180
QY 181 CTCTGGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
DB 181 CTCTGGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
QY 241 TCGGGATGTACCATGTCAAGACGACTGCTCAACTCAAGCATTTGTTATGAGGAGCG 300
DB 241 TCGGGATGTACCATGTCAAGACGACTGCTCAACTCAAGCATTTGTTATGAGGAGCG 300
QY 301 GACATGATCATGACACACCCCGGGTGGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
DB 301 GACATGATCATGACACACCCCGGGTGGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
QY 361 TGCTGGGTAGCGCTACACCCCGGGTGGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
DB 361 TGCTGGGTAGCGCTACACCCCGGGTGGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
QY 421 ATACGACGCGCAGCTGCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
DB 421 ATACGACGCGCAGCTGCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
QY 481 GGGGACCTCTGCGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
DB 481 GGGGACCTCTGCGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
QY 541 CATGACGCGTGCAGGACTGCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
DB 541 CATGACGCGTGCAGGACTGCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
QY 601 ATGGCTTGGGATATGATGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
DB 601 ATGGCTTGGGATATGATGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
QY 661 CTCGGATCCCAAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
DB 661 CTCGGATCCCAAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720

Db	541		CATGAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGGTCCACCGT	600
Qy	601	ATGSGCTTTGGGATATGATGAACTGGTTCGCTACAAACGGGCCCTGGTGGTATCGAGCTG	660	
Db	601	ATGSGCTTTGGGATATGATGAACTGGTTCGCTACAAACGGGCCCTGGTGGTATCGAGCTG	660	
Qy	661	CTCGGATCCCAACAGCTGCTGGACAATGTTGGCGGGGGGCCCAATCGGAGAGTCTCTGGCG	720	
Db	661	CTCGGATCCCAACAGCTGCTGGACAATGTTGGCGGGGGGCCCAATCGGAGAGTCTCTGGCG	720	
Qy	721	GGTCTCGCCTACTATTCCATCGTGGGGGAACTGGGCTAAGCTTTTGTATGTTGATGCTACTC	780	
Db	721	GGTCTCGCCTACTATTCCATCGTGGGGGAACTGGGCTAAGCTTTTGTATGTTGATGCTACTC	780	
Qy	781	TTTGCTCCCTAATAG	795	
Db	781	TTTGCTCCCTAATAG	795	
RESULT 3				
LOCUS	AX452754			
DEFINITION	Sequence 5 from Patent EPI211315.	795 bp	DNA	linear
ACCESSION	AX452754			
VERSION	AX452754.1	GI:21712439		
KEYWORDS				
SOURCE	Hepatitis C virus			
ORGANISM	Hepatitis C virus			
REFERENCE				
AUTHORS	Maertens,G., Bosman,F., de Martynoff,G. and Buysse,M.A.			
TITLE	Recombinant vectors for producing hcv envelope proteins			
JOURNAL	Patent: EP 1211315-A 5 05-JUN-2002;			
INNOGENETICS	Innogenetics N.V. (BE)			
FEATURES				
source		Location/Qualifiers		
		1..795		
		/organism="Hepatitis C virus"		
		/mol_type="unassigned DNA"		
		/db_xref="taxon:11103"		
CDS		1..792		
		/note="unnamed protein product"		
		/codon_start=1		
		/protein_id="CAD38067.1"		
		/db_xref="GI:21712440"		
		/db_xref="REMBL:CAD38067"		
		/translation="MLGKVIDTLTCGFADLVGYIPLVGAPLGGARALAHGVRLSDG		
		VNTATMLPGCSIFILLALUSCLTPASAEYVRNVSGMYHTVNDSCNSISVVEADDM		
		IMHTPGCVPCVRNNSSRCRVALTPTLAAENASVPTTIRHVDLVLGAAAFCSAMY		
		GDLCFVFLVSQLTFSRPHETVDCNCISYIPGHITGHRMADMMNNMSPPTTALVVS		
		QLLRIPQAVDMVAGAHGVLGLAYYSVMGNKAKVLIVMLLPAP"		
mat_peptide		1..789		
		/product="unnamed"		
ORIGIN				
Query Match		100.0%;	Score 795; DB 6;	Length 795;
Best Local Similarity		100.0%;	Pred. No. 3e-167;	
Matches 795, Conservative		0; Mismatches	0; Indels	0; Gaps
Qy	1	ATGTTGGGTAAAGTTCATGATACCTTTACATGCGGCTTCGCCGACCTCGTGGGGTACATT	60	
Db	1	ATGTTGGGTAAAGTTCATGATACCTTTACATGCGGCTTCGCCGACCTCGTGGGGTACATT	60	
Qy	61	CCGCTCGTCGGCGCCCTTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGGTCGGGTT	120	
Db	61	CCGCTCGTCGGCGCCCTTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGGTCGGGTT	120	
Qy	121	CTGAGGACCGCGTGAATATGCAACAGGGAATTTGCCCGGTGCTCTTTCTCTATCTTC	180	
Db	121	CTGAGGACCGCGTGAATATGCAACAGGGAATTTGCCCGGTGCTCTTTCTCTATCTTC	180	
Qy	181	CTCTTGGCTTTGCTGCTGCTTGACCCGTTCCAGCTTCCGCTATGAGTGCACAGCTG	240	

VNYATGNLPGCSFSPFIPLALLSCLTVPASAVEVRNCSGWHVTDNCSSNIVYEADAM
 IMHTPGCVCPRENNSSRCWVALPTPLAARNASVPTTIRRHVLDVGAFAFCSAMTV
 DGLGSGVPLVSQLOFTISPRRHETVQDCNCSIPGHIITCHRNAWDMANWSPETALVVS
 QLLRIPOAVDVMVAGAHGVLGLAGLAYTISVGVGNKAKVLIVMLLFP*
 1. .789
 /product="unnamed"

mat peptide

1.789

```
product="unnamed"
```

ORIGIN

Query Match	100 Q&S: Score 795:	DB 6: Length 795:
1	100 Q&S: Score 795:	DB 6: Length 795:

Query Match	100 Q&S: Score 795:	DB 6: Length 795:
1	100 Q&S: Score 795:	DB 6: Length 795:

Query Match	100.00%	Score	755	DB #	155
Best Local Similarity	100.0%	Pred. No.	3e-167		
Matches	795	Mismatches	0	Indels	0
Matches	795	Conservative	0	Gaps	0

ON 1 ATGTTTGGCTAAGCTTCAATCGATACCTTACATGCGCTTCCGACGACCTGCGCTACATT 60

1 A T C G T T C C G C T A A G C C T C A T C G A T A C C C T T A C A T C C G C G C C G T A C A T T 60

[illegible]

120

121

122

123

124

125

126

127

128

129

130

131

132

133

134

135

136

137

138

139

140

141

142

143

144

145

146

147

148

149

150

151

152

153

154

155

156

157

158

159

160

161

162

163

164

165

166

167

168

169

170

171

172

173

174

175

176

177

178

179

180

181

182

183

184

185

186

187

188

189

190

191

192

193

194

195

196

197

198

199

200

201

202

203

204

205

206

207

208

209

210

211

212

213

214

215

216

217

218

219

220

221

222

223

224

225

226

227

228

229

230

231

232

233

234

235

236

237

238

239

240

241

242

243

244

245

246

247

248

249

250

251

252

253

254

255

256

257

258

259

260

261

262

263

264

265

266

267

268

269

270

271

272

273

274

275

276

277

278

279

280

281

282

283

284

285

286

287

288

289

290

291

292

293

294

295

296

297

298

299

300

301

302

303

304

305

306

307

308

309

310

311

312

313

314

315

316

317

318

319

320

321

322

323

324

325

326

327

328

329

330

331

332

333

334

335

336

337

338

339

340

341

342

343

344

345

346

347

348

349

350

351

352

353

354

355

356

357

358

359

360

361

362

363

364

365

366

367

368

369

370

371

372

373

374

375

376

377

378

379

380

381

382

383

384

385

386

387

388

389

390

391

392

393

394

395

396

397

398

399

400

401

402

403

404

405

406

407

408

409

410

411

412

413

414

415

416

417

418

419

420

421

422

423

424

425

426

427

428

429

430

431

432

433

434

435

436

437

438

439

440

441

442

443

444

445

446

447

448

449

450

451

452

453

454

455

456

457

458

459

460

461

462

463

464

465

466

467

468

469

470

471

472

473

474

475

476

477

478

479

480

481

482

483

484

485

486

487

488

489

490

491

492

493

494

495

496

497

498

499

500

501

502

503

504

505

506

507

508

509

510

511

512

513

514

515

516

517

518

519

520

521

522

523

524

525

526

527

528

529

530

531

532

533

534

535

536

537

538

539

540

541

542

543

544

545

546

547

548

549

550

551

552

553

554

555

556

557

558

559

560

561

562

563

564

565

566

567

568

569

570

571

572

573

574

575

576

577

578

579

580

581

582

583

584

585

586

587

588

589

590

591

592

593

594

595

596

597

598

599

600

601

602

603

604

605

606

607

608

609

610

611

612

613

614

615

616

617

618

619

620

621

622

623

624

625

626

627

628

629

630

631

<

.....

97

DD 1 CCGGKGTGTACCAATGTCACGATCGATCGTCCATCTCAAGCAATGTGTATGTAGGAGAGCG 300

QY 301 GHAAGATCAAGCACCCCGGGTGGTGCCTGGGTGGGAGAACATCTCCCGC 380

DD 301 GACATGATCAATGCACACCCCGGGTGGTGGTGGGAGAACAACTCTTCCGC 360

QY 361 TGC TGGGTAGCGCTCACCCCAAGCTCGCAGCTAGGAACGCCAGCGTCCCCACCGACA 420

D_b 361 TGCTGGGTAGCGCTCACCCACGCTCGCAGCTAGGAACGCCAGCGTCCCCACCGACA 420

Qy 421 ATACGACGCCACGTCGATTTCGTTCGTCGGCGGCTGCTTTCGTTCGGCTATGTACGTG 480

421 ATACGACGCCACGTGATTTCCTCGTTGGGGCGGCTCTTCTGTTCGGCTATGTACGTG 480

DV 481 GGGGACCTATTCGCGGATTCTGTCTCCAGTGTCAACAATTACGCACCG 540

100

2000

A

QY /21 GGTCGCGCACATATCCATGGTGGGGAACGGGCATAGGTATTGATATGATGCTACTC 780

DB /21 GGCTCGCCTACTATTCACAGGTGGGGAAC TGGCTAAGGTTTGATGTGATGCTACAC 780

PRISTIT 5

AY791819

10010
10010

SUNNY
89486/VH
COMMENCE FROM DUTCH W003051012
DNA
TINERAT FBI 04-OCT-2000

795 40

אנדרסן

LINEAR DATA-2003


```
FEATURES             source
1..2082
/organism="Hepatitis C virus"
/mol_type="unassigned DNA"
/db_xref="taxon:11103"
<1..2079
/feature="unassigned protein product"
/codon_start=1
/db_xref="GI:37515780"
/translation="NLGKVIDTLTCGPFADLVGYIPLVGAIPGAGARALAHGVRLVLEGG
VNYATGNLPGCSFISIFLLALLSLCTVPASAYEVNRVSMYHVTNDCSNSSIVYEAADM
IMHTPGCVCRVRENNSSRCWALPTLAARNASVPTTIRRHVDLLVGAAPCSAMY
GDLCSFVLVSQFLPTISPRHETVQDCNSIYPGHI TGRHMAWMMNWSPTTALVVS
QLLRIPQAVDMVAGAHGVLGLAYYSMVGNAKVLVVMLLPAGVDCHTRVSGCAA
SDTRGLVSLFSPGSAQKIQLVNTGSHWINTALNCNDSLOTGTPFAALPYKHKNSSG
CPEPLASCRSIDKFAQMGWGLTYTEPNSSDQPCYCMHYAPRPGGIVPASQVGPVYCF
TPSPVVGTTDRFQVYNNGANDSDVLILANTPRPGNWFCTGMMNGTGPTKCGGP
PCNTGGAGNNTLTCTDCFRKHPEATYARCSGPMWTPRCMVHYPTVRLMHPYCTVNF
IPKVMYGVGVEHREFAACNWRGRCDEDRSELSPLLSITTEMWILLPCSTPTLP
ALSTGLIHLHONIVDVOYLGVGSVAVSLVTKWEYVLLFLILLADARICACIMMMLL
AQAEALLENLVLMRAAVAGAHGTLFLVFFCAAWYIKGRIVPGNAIYAFYGVWPLLL
LLALUPPPRAYA"
mat_peptide
1..2076
/product="unnamed"

ORIGIN
Query Match      98.0%; Score 778.8; DB 6; Length 2082;
Best Local Similarity 99.7%; Pred. No. 1.2e-163;
Matches 780; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 TTGGTAAAGTTCATGATACCTTACATCGCGCTTCGCGACCTCGTGGGGTACATTCG 63
DB 4 TTGGTAAAGTTCATGATACCTTACATCGCGCTTCGCGACCTCGTGGGGTACATTCG 63

QY 64 CTGCTCGCGCGCCCCCTAGAGGGGCGCTGCGAGGGCCCTGCGGCATCGCGTTCG 123
DB 64 CTGCTCGCGCGCCCCCTAGAGGGGCGCTGCGAGGGCCCTGCGGCATCGCGTTCG 123

QY 124 GAGGACGGGTGAACATATGCAACAGGGAATTTGCCGGTGTCTTCTCTATCTTCCTC 183
DB 124 GAGGACGGGTGAACATATGCAACAGGGAATTTGCCGGTGTCTTCTCTATCTTCCTC 183

QY 184 TTGGCTTTGCTGCTCTGCTGACCTTCAGCTTCGCTTCCAGCTTGAAGTGGCAGCGTCC 243
DB 184 TTGGCTTTGCTGCTCTGCTGACCTTCAGCTTCGCTTCCAGCTTGAAGTGGCAGCGTCC 243

QY 244 GGGATGTACCATGTCAAGAACGACTGCTCAACTCAAGCATTTGTATGAGGCGAGCGAC 303
DB 244 GGGATGTACCATGTCAAGAACGACTGCTCAACTCAAGCATTTGTATGAGGCGAGCGAC 303

QY 304 ATGATCATGCACACCCCGGGTGGTGGCTCGCTCGGTTGGGAGAAACAACTCTCCGCTGC 363
DB 304 ATGATCATGCACACCCCGGGTGGTGGCTCGCTCGGTTGGGAGAAACAACTCTCCGCTGC 363

QY 364 TGGGTAGGCTTACCCCGACGCTCGCAGTGAAGACCGCAGCGTCCACACGACAATA 423
DB 364 TGGGTAGGCTTACCCCGACGCTCGCAGTGAAGACCGCAGCGTCCACACGACAATA 423

QY 424 CGACGCCACGCTGATTTGCTGTGGGGGGCTGCTTCTGTTCCGCTATGTACGTGGGG 483
DB 424 CGACGCCACGCTGATTTGCTGTGGGGGGCTGCTTCTGTTCCGCTATGTACGTGGGG 483

QY 484 GACCTCTCGGATCTGTCTTCTCTGCTCCAGCTGTTTCCACCATCTCGCTCCCGGCAT 543
DB 484 GACCTCTCGGATCTGTCTTCTCTGCTCCAGCTGTTTCCACCATCTCGCTCCCGGCAT 543

QY 544 GAGACGGTGCAGGATGCTCAATCTATATCCCGGCCACATTAACGGGTCAACGTTATG 603
DB 544 GAGACGGTGCAGGATGCTCAATCTATATCCCGGCCACATTAACGGGTCAACGTTATG 603

QY 604 GCTTCGGATATCATGATGATGCTGCTTACAAAGCGCCCTGCTGTATCGGAGTCTC 663
DB 604 GCTTCGGATATCATGATGATGCTGCTTACAAAGCGCCCTGCTGTATCGGAGTCTC 663

FEATURES             source
1..2433
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"
1..2430
/feature="unnamed protein product"
/codon_start=1
/protein_id="CAA03147.1"
/db_xref="GI:2302425"
/translation="NSTWPKPQKTRNTNRPPQDVKPPGGQIVGVVYLLPRRGRPL
GVRAATKTSERSQPRGRQPIKARPEGRWAQPCYDWPVLYGNCMGWAGMLLSPRG
SRPSWGTDDPRRNRNLGKVIDTLTCGPFADLVGYIPLVGAIPGAGARALAHGVRLV
GVNYATGNLPGCSFISIFLLALLSLCTVPASAYEVNRVSMYHVTNDCSNSSIVYEAAD
MIMHTPGCVCRVRENNSSRCWALPTLAARNASVPTTIRRHVDLLVGAAPCSAMY
VDLCSFVLVSQFLPTISPRHETVQDCNSIYPGHI TGRHMAWMMNWSPTTALVVS
SOLLRIPOAVDMVAGAHGVLGLAYYSMVGNAKVLVVMLLPAGVDCHTRVSGCAA
ASDTRGLVSLFSPGSAQKIQLVNTGSHWINTALNCNDSLOTGTPFAALPYKHKNSS
GCPERLASCRSIDKFAQMGWGLTYTEPNSSDQPCYCMHYAPRPGGIVPASQVGPVYCF
TPSPVVGTTDRFQVYNNGANDSDVLILANTPRPGNWFCTGMMNGTGPTKCGGP
PCNTGGAGNNTLTCTDCFRKHPEATYARCSGPMWTPRCMVHYPTVRLMHPYCTVNF
IPKVMYGVGVEHREFAACNWRGRCDEDRSELSPLLSITTEMWILLPCSTPTLP
ALSTGLIHLHONIVDVOYLGVGSVAVSLVTKWEYVLLFLILLADARICACIMMMLL
AQAEALLENLVLMRAAVAGAHGTLFLVFFCAAWYIKGRIVPGNAIYAFYGVWPLLL
LLALUPPPRAYA"
mat_peptide
1..2427
/product="unnamed"

ORIGIN
Query Match      98.0%; Score 778.8; DB 6; Length 2433;
Best Local Similarity 99.7%; Pred. No. 1.2e-163;
Matches 780; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 TTGGTAAAGTTCATGATACCTTACATCGCGCTTCGCGACCTCGTGGGGTACATTCG 63
DB 355 TTGGTAAAGTTCATGATACCTTACATCGCGCTTCGCGACCTCGTGGGGTACATTCG 414

QY 64 CTGCTCGCGCGCCCCCTAGAGGGGCGCTGCGAGGGCCCTGCGGCATCGCGTTCG 123
DB 64 CTGCTCGCGCGCCCCCTAGAGGGGCGCTGCGAGGGCCCTGCGGCATCGCGTTCG 123
```



```
Db 415 CTGCTGGGCGCCCTAGGGGGCGCTGCCAGGGCCCTGGGCGATGGGGTCCGGGTTCTG 474
Qy 124 GAGGACGGCGTGAACATATGCAACAGGGAATTTGCCGGTGTCTCTTTCTATCTTCTC 183
Db 475 GAGGACGGCGTGAACATATGCAACAGGGAATTTGCCGGTGTCTCTTTCTATCTTCTC 534
Qy 184 TTGGCTTTGCTGCTCTGCTGACCGTTCCAGCTTCCGCTTATGAAGTGGCAAGTGTCC 243
Db 535 TTGGCTTTGCTGCTCTGCTGACCGTTCCAGCTTCCGCTTATGAAGTGGCAAGTGTCC 594
Qy 244 GCGATGTACCATGTGACGAAAGCATGTCTCAACTCAAGCATTTGTATGAGGCGAGCGAC 303
Db 595 GCGATGTACCATGTGACGAAAGCATGTCTCAACTCAAGCATTTGTATGAGGCGAGCGAC 654
Qy 304 ATGATCATGACACACCCCGGGTGTGCTGCTCCGCTTCCGGAGAAACAATCTTCCGCTGC 363
Db 655 ATGATCATGACACACCCCGGGTGTGCTGCTCCGCTTCCGGAGAAACAATCTTCCGCTGC 714
Qy 364 TGGGTAGCGCTCACCCCGGCTGCTGAGGAGCGCGCTCTTCTGCTGCTATGATGATGAGG 423
Db 715 TGGGTAGCGCTCACCCCGGCTGCTGAGGAGCGCGCTCTTCTGCTGCTATGATGATGAGG 774
Qy 424 CGAGCGACAGTGCATTTGCTGCTGAGGCGCGCTCTTCTGCTGCTATGATGATGAGG 483
Db 775 CGAGCGACAGTGCATTTGCTGCTGAGGCGCGCTCTTCTGCTGCTATGATGATGAGG 834
Qy 484 GACCTCTGCGGATCTGTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 543
Db 835 GACCTCTGCGGATCTGTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 894
Qy 544 GAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCGACATAAAGCGGTACCGTATG 603
Db 895 GAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCGACATAAAGCGGTACCGTATG 954
Qy 604 GCTTGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 663
Db 955 GCTTGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1014
Qy 784 GC 785
Db 1135 GC 1136
```

```
RESULT 12
LOCUS AR157351 2433 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 49 from patent US 6245503.
ACCESSION AR157351
VERSION AR157351.1 GI:16218285
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2433)
AUTHORS Maertens, G., Bosman, F., De Martynoff, G. and Buyse, M.-A.
TITLE Purified hepatitis C virus envelope proteins for diagnostic and therapeutic use
JOURNAL Patent: US 6245503-A 49 12-JUN-2001;
FEATURES Location/Qualifiers
source 1..2433
/mol_type="unassigned DNA"
```

Query Match 98.0%; Score 778.8; DB 6; Length 2433;

```
Best Local Similarity 99.7%; Pred. No. 1.2e-163;
Matches 780; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 4 TTGGGTAAAGGTGATCATGATACCTTACATGCGGCTTTCGCGACTCGTGGGGTACATTCGG 63
Db 355 TTGGGTAAAGGTGATCATGATACCTTACATGCGGCTTTCGCGACTCGTGGGGTACATTCGG 414
Qy 64 CTGCTGCGGCGCCCTAGAGGGGCGCTGCCAGGGCCCTGGGCGCATGGGGTTCG 123
Db 415 CTGCTGCGGCGCCCTAGAGGGGCGCTGCCAGGGCCCTGGGCGCATGGGGTTCG 474
Qy 124 GAGGACGGCGTGAACATATGCAACAGGGAATTTGCCGGTGTCTCTTTCTATCTTCTC 183
Db 475 GAGGACGGCGTGAACATATGCAACAGGGAATTTGCCGGTGTCTCTTTCTATCTTCTC 534
Qy 184 TTGGCTTTGCTGCTCTGCTGACCGTTCCAGCTTCCGCTTATGAAGTGGCAAGTGTCC 243
Db 535 TTGGCTTTGCTGCTCTGCTGACCGTTCCAGCTTCCGCTTATGAAGTGGCAAGTGTCC 594
Qy 244 GCGATGTACCATGTGACGAAAGCATGTCTCAACTCAAGCATTTGTATGAGGCGAGCGAC 303
Db 595 GCGATGTACCATGTGACGAAAGCATGTCTCAACTCAAGCATTTGTATGAGGCGAGCGAC 654
Qy 304 ATGATCATGACACACCCCGGGTGTGCTGCTCCGCTTCCGGAGAAACAATCTTCCGCTGC 363
Db 655 ATGATCATGACACACCCCGGGTGTGCTGCTCCGCTTCCGGAGAAACAATCTTCCGCTGC 714
Qy 364 TGGGTAGCGCTCACCCCGGCTGCTGAGGAGCGCGCTCTTCTGCTGCTATGATGATGAGG 423
Db 715 TGGGTAGCGCTCACCCCGGCTGCTGAGGAGCGCGCTCTTCTGCTGCTATGATGATGAGG 774
Qy 424 CGAGCGACAGTGCATTTGCTGCTGAGGCGCGCTCTTCTGCTGCTATGATGATGAGG 483
Db 775 CGAGCGACAGTGCATTTGCTGCTGAGGCGCGCTCTTCTGCTGCTATGATGATGAGG 834
Qy 484 GACCTCTGCGGATCTGTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 543
Db 835 GACCTCTGCGGATCTGTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 894
Qy 544 GAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCGACATAAAGCGGTACCGTATG 603
Db 895 GAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCGACATAAAGCGGTACCGTATG 954
Qy 604 GCTTGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 663
Db 955 GCTTGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1014
Qy 664 CGGATCCCAAGCTGTGCTGCGCATGCTGCGGCGGCGCCATTTGGGGAGTCTTGGCGGGT 723
Db 1015 CGGATCCCAAGCTGTGCTGCGCATGCTGCGGCGGCGCCATTTGGGGAGTCTTGGCGGGC 1074
Qy 724 CTGCGCTACTATTCCATGTTGGGGAACCTGCGCTACAAAGCGCCCTGGTGGTATGCGAGCTGCTC 783
Db 1075 CTGCGCTACTATTCCATGTTGGGGAACCTGCGCTACAAAGCGCCCTGGTGGTATGCGAGCTGCTC 1134
Qy 784 GC 785
Db 1135 GC 1136
```

```
RESULT 13
LOCUS AX452798 2433 bp DNA linear PAT 06-JUL-2002
DEFINITION Sequence 49 from Patent EP1211315.
ACCESSION AX452798
VERSION AX452798.1 GI:21712483
KEYWORDS
SOURCE Hepatitis C virus
ORGANISM Hepatitis C virus
REFERENCE 1
AUTHORS Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
Maertens, G., Bosman, F., de Martynoff, G. and Buyse, M.-A.
```

TITLE

JOURNAL
Patent: EP 1211315-A 49 05-JUN-2002:

InnoGenetics N.V. (BE)

FEATURES	Location/Qualifiers
1. <i>the</i>	1. <i>the</i> (100%)
2. <i>the</i>	2. <i>the</i> (100%)
3. <i>the</i>	3. <i>the</i> (100%)
4. <i>the</i>	4. <i>the</i> (100%)
5. <i>the</i>	5. <i>the</i> (100%)
6. <i>the</i>	6. <i>the</i> (100%)
7. <i>the</i>	7. <i>the</i> (100%)
8. <i>the</i>	8. <i>the</i> (100%)
9. <i>the</i>	9. <i>the</i> (100%)
10. <i>the</i>	10. <i>the</i> (100%)
11. <i>the</i>	11. <i>the</i> (100%)
12. <i>the</i>	12. <i>the</i> (100%)
13. <i>the</i>	13. <i>the</i> (100%)
14. <i>the</i>	14. <i>the</i> (100%)
15. <i>the</i>	15. <i>the</i> (100%)
16. <i>the</i>	16. <i>the</i> (100%)
17. <i>the</i>	17. <i>the</i> (100%)
18. <i>the</i>	18. <i>the</i> (100%)
19. <i>the</i>	19. <i>the</i> (100%)
20. <i>the</i>	20. <i>the</i> (100%)
21. <i>the</i>	21. <i>the</i> (100%)
22. <i>the</i>	22. <i>the</i> (100%)
23. <i>the</i>	23. <i>the</i> (100%)
24. <i>the</i>	24. <i>the</i> (100%)
25. <i>the</i>	25. <i>the</i> (100%)
26. <i>the</i>	26. <i>the</i> (100%)
27. <i>the</i>	27. <i>the</i> (100%)
28. <i>the</i>	28. <i>the</i> (100%)
29. <i>the</i>	29. <i>the</i> (100%)
30. <i>the</i>	30. <i>the</i> (100%)
31. <i>the</i>	31. <i>the</i> (100%)
32. <i>the</i>	32. <i>the</i> (100%)
33. <i>the</i>	33. <i>the</i> (100%)
34. <i>the</i>	34. <i>the</i> (100%)
35. <i>the</i>	35. <i>the</i> (100%)
36. <i>the</i>	36. <i>the</i> (100%)
37. <i>the</i>	37. <i>the</i> (100%)
38. <i>the</i>	38. <i>the</i> (100%)
39. <i>the</i>	39. <i>the</i> (100%)
40. <i>the</i>	40. <i>the</i> (100%)
41. <i>the</i>	41. <i>the</i> (100%)
42. <i>the</i>	42. <i>the</i> (100%)
43. <i>the</i>	43. <i>the</i> (100%)
44. <i>the</i>	44. <i>the</i> (100%)
45. <i>the</i>	45. <i>the</i> (100%)
46. <i>the</i>	46. <i>the</i> (100%)
47. <i>the</i>	47. <i>the</i> (100%)
48. <i>the</i>	48. <i>the</i> (100%)
49. <i>the</i>	49. <i>the</i> (100%)
50. <i>the</i>	50. <i>the</i> (100%)
51. <i>the</i>	51. <i>the</i> (100%)
52. <i>the</i>	52. <i>the</i> (100%)
53. <i>the</i>	53. <i>the</i> (100%)
54. <i>the</i>	54. <i>the</i> (100%)
55. <i>the</i>	55. <i>the</i> (100%)
56. <i>the</i>	56. <i>the</i> (100%)
57. <i>the</i>	57. <i>the</i> (100%)
58. <i>the</i>	58. <i>the</i> (100%)
59. <i>the</i>	59. <i>the</i> (100%)
60. <i>the</i>	60. <i>the</i> (100%)
61. <i>the</i>	61. <i>the</i> (100%)
62. <i>the</i>	62. <i>the</i> (100%)
63. <i>the</i>	63. <i>the</i> (100%)
64. <i>the</i>	64. <i>the</i> (100%)
65. <i>the</i>	65. <i>the</i> (100%)
66. <i>the</i>	66. <i>the</i> (100%)
67. <i>the</i>	67. <i>the</i> (100%)
68. <i>the</i>	68. <i>the</i> (100%)
69. <i>the</i>	69. <i>the</i> (100%)
70. <i>the</i>	70. <i>the</i> (100%)
71. <i>the</i>	71. <i>the</i> (100%)
72. <i>the</i>	72. <i>the</i> (100%)
73. <i>the</i>	73. <i>the</i> (100%)
74. <i>the</i>	74. <i>the</i> (100%)
75. <i>the</i>	75. <i>the</i> (100%)
76. <i>the</i>	76. <i>the</i> (100%)
77. <i>the</i>	77. <i>the</i> (100%)
78. <i>the</i>	78. <i>the</i> (100%)
79. <i>the</i>	79. <i>the</i> (100%)
80. <i>the</i>	80. <i>the</i> (100%)
81. <i>the</i>	81. <i>the</i> (100%)
82. <i>the</i>	82. <i>the</i> (100%)
83. <i>the</i>	83. <i>the</i> (100%)
84. <i>the</i>	84. <i>the</i> (100%)
85. <i>the</i>	85. <i>the</i> (100%)
86. <i>the</i>	86. <i>the</i> (100%)
87. <i>the</i>	87. <i>the</i> (100%)
88. <i>the</i>	88. <i>the</i> (100%)
89. <i>the</i>	89. <i>the</i> (100%)
90. <i>the</i>	90. <i>the</i> (100%)
91. <i>the</i>	91. <i>the</i> (100%)
92. <i>the</i>	92. <i>the</i> (100%)
93. <i>the</i>	93. <i>the</i> (100%)
94. <i>the</i>	94. <i>the</i> (100%)
95. <i>the</i>	95. <i>the</i> (

Source

/organism="Hepatitis C virus"

```
/mol_type="unassigned DNA"
```

```
/db_xref="taxon:11103"
```

1. 2430

/note="unnamed protein produc

```
/codon_start=1
```

/protein_id="CAD38083.1"

```

/db_xref="GI:21712484"

```

```
/db_xref="REMFREMBL:CAD38083"
```

```
/translation="MS-INP-KP-QR-KT-KR-NI-
```

GVAZKATLSEKSGRGRKQV1DKLKPGRGARAUGQ1PFLTJLGGARAGWALGVRLD
 RSPWSDPTLRRSJRNKQV1DTLTCGFADLVGY1VLGPALGGARAGWALGVRLD
 GUNYATGLPCPFS1FLLALLSLTLPASVYEVRSVHTVNDLCSNS1VYRAAD
 M1MHTPGCPGCRNNSSCWALVPTLAAARNASVPTT1IRRHVTLVGCAGAFCSAA
 VGLD-CGSEVLVSQ1FLT1SPRHEVTQDCNS1LPCGH1TGRHMAWMMWNSPTTALV
 SOLLRT1QA1VDMVGAHGWGLAGLY1YSVGNWAKV1VWML1PAGVDCGHTVSGAA
 GTPERLAS1CS1DKRPAQWGL1TTPENS1SDQRY1CMHWAP1CG1VPASVQGVPC
 GFCPS1VVTG1DTFGVPTVYNDSS1DLV1LNT1P1RG1CGT1WNGG1GFTK1CGG
 PPN1IGAGNN1L1AC1PDC1EK1HP1EAT1VARG1SG1PWL1PRCMV1P1RL1WHP1CTV1F
 T1PKV1MY1GVSEH1P1V01L1GVGS1AV1SLV1TKW1EVL1LL1FLL1ADAR1I1CACL1WMLL1
 PALST1GL1HL1HQN1V1D1Y1L1GVGS1AV1SLV1TKW1EVL1LL1FLL1ADAR1I1CACL1WMLL1
 1IAQEAAL1EN1VL1N1AA1VAG1H1TL1FLV1FF1CA1W1Y1KGR1LV1P1G1AA1Y1FY1GV1W1PLL1
 1LL1AP1PR1VA1"

1. .2427

```
/product="unnamed"
```

ORIGIN

Query Match	98.0%;	Score 778.8;	DB 6;	Length 2433;
Best Local Similarity	99.7%;	Pred. No. 1.2e-163;		
Matches 780;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;
4	TTGGGTAAAGGTCAATCGATACCCCTTACATGCGCGCTTGGCGGACCTCGTGGGGTACATTCGG	63		
355	TTGGGTAAAGGTCAATCGATACCCCTTACATGCGCGCTTGGCGGACCTCGTGGGGTACATTCGG	414		
64	CTCGTCGGCGCCCCCTTAGGGGGCGCTGCCAGGGGCCCTGGCGGCAATGGCGTTCGGGTTCTG	123		
415	CTCGTCGGCGCCCCCTTAGGGGGCGCTGCCAGGGGCCCTGGCGGCAATGGCGTTCGGGTTCTG	474		
124	GAGGACGGCGTGAACTATGCAACAGGGGAATTTGCCCGGTGCTCTTTCTCTATCTTCCTC	183		
475	GAGGACGGCGTGAACTATGCAACAGGGGAATTTGCCCGGTGCTCTTTCTCTATCTTCCTC	534		
184	TTGGGCTTTTGCTGCTCTGCTGACCGTTCCAGCTTCGGCTATGAAGTGCAGCAAGTGTCC	243		
535	TTGGGCTTTTGCTGCTCTGCTGACCGTTCCAGCTTCGGCTATGAAGTGCAGCAAGTGTCC	594		
244	GGGATGTACCATGTCACGAAACGACTGCTCCAACTCAAGCATTTGTATGAGCGACGGAC	303		
595	GGGATGTACCATGTCACGAAACGACTGCTCCAACTCAAGCATTTGTATGAGCGACGGAC	654		
304	ATGATCATGCAACACCCCGGGTGGTGGCCCTGGCGTTGGGAGAACACTCTTCCCGCTGC	363		
655	ATGATCATGCAACACCCCGGGTGGTGGCCCTGGCGTTGGGAGAACACTCTTCCCGCTGC	714		
364	TGGGTAGCGCTCACCCCGCAACGCTCCGACGTAGGAACGCCAGGCTCCCAACACCAACAATA	423		
715	TGGGTAGCGCTCACCCCGCAACGCTCCGACGTAGGAACGCCAGGCTCCCAACACCAACAATA	774		
424	CGACGCCACGTGCAATTTGCTCGTTGGGGCGGCTGCTTTCTGTTCCGCTATGTAGCTGGGG	483		
775	CGACGCCACGTGCAATTTGCTCGTTGGGGCGGCTGCTTTCTGTTCCGCTATGTAGCTGGGG	834		
484	GACCTCTGCGGATCTGTCTTCTCCGCTCTCCACGCTGTTTCAACATCTGCGCTGCCGGGAT	543		
835	GACCTCTGCGGATCTGTCTTCTCCGCTCTCCACGCTGTTTCAACATCTGCGCTGCCGGGAT	894		

Db	355	TTGGGTAAAGTTCATCGATACCTTTACATGCGGCTTCGCGACCTCGTGGGTACATCCG	414
Qy	64	CTCGTGGGCGCCCTTAGGGGGCGCTGCCAGGGCCCTTGGCGCATGCGTCCGGTTCG	123
Db	415	CTCGTGGGCGCCCTTAGGGGGCGCTGCCAGGGCCCTTGGCGCATGCGTCCGGTTCG	474
Qy	124	GAGGAGCGGTGAACATATGCAACAGGGAATTTGCGCGGTTCCTTTCTCTATCTTCTC	183
Db	475	GAGGAGCGGTGAACATATGCAACAGGGAATTTGCGCGGTTCCTTTCTCTATCTTCTC	534
Qy	184	TTGGCTTTGCTGCTCTGACCGTTCCAGCTTCCGCTTATGAGTGGCAAGCTGTCC	243
Db	535	TTGGCTTTGCTGCTCTGACCGTTCCAGCTTCCGCTTATGAGTGGCAAGCTGTCC	594
Qy	244	GCGATGTACCATGTTCAGAACGACTGCTCCAACTCAAGCATTTGTATGAGGACAGCGAC	303
Db	595	GCGATGTACCATGTTCAGAACGACTGCTCCAACTCAAGCATTTGTATGAGGACAGCGAC	654
Qy	304	ATGATCATGACACCCCGGGTGGTGGCTTCCGCTTCCGAGAACAACTCTTCCCGTGC	363
Db	655	ATGATCATGACACCCCGGGTGGTGGCTTCCGCTTCCGAGAACAACTCTTCCCGTGC	714
Qy	364	TGGGTAGCGCTCACCCGCCACGCTGCGAGTGGCAAGCGCAGCGTCCGCCACACAAATA	423
Db	715	TGGGTAGCGCTCACCCGCCACGCTGCGAGTGGCAAGCGCAGCGTCCGCCACACAAATA	774
Qy	424	CGACGCCACGTCGATTTGCTGTTGGGGCGGCTGCTTTCTGCTGCTATGTACGTGGG	483
Db	775	CGACGCCACGTCGATTTGCTGTTGGGGCGGCTGCTTTCTGCTGCTATGTACGTGGG	834
Qy	484	GACCTCTGCGGATCTGCTTCTCGTCTCCAGCTGTTTCCACATCTCGCTCGCGCAT	543
Db	835	GACCTCTGCGGATCTGCTTCTCGTCTCCAGCTGTTTCCACATCTCGCTCGCGCAT	894
Qy	544	GAGCGGTGACGAGTGCATTTGCTCAATCTATCCCGGCCACATPACGGGTTCACCGTATG	603
Db	895	GAGCGGTGACGAGTGCATTTGCTCAATCTATCCCGGCCACATPACGGGTTCACCGTATG	954
Qy	604	GCTTGGATATGATGATGACTGCTGCTTGGGGCGGCTGCTTCTGCTGCTATGTACGTGGG	663
Db	955	GCTTGGATATGATGATGACTGCTGCTTGGGGCGGCTGCTTCTGCTGCTATGTACGTGGG	1014
Qy	664	CGGATCCCAAGCTGCTGCGGACATGCTGGCGGGGGCCCATTTGGGAGTCTCTGCGGGT	723
Db	1015	CGGATCCCAAGCTGCTGCGGACATGCTGGCGGGGGCCCATTTGGGAGTCTCTGCGGGC	1074
Qy	724	CTCGCTTACTTATCCATGGTGGGAACTGGGCTTAAGTTTGTATGATGCTACTCTTT	783
Db	1075	CTCGCTTACTTATCCATGGTGGGAACTGGGCTTAAGTTTGTATGATGCTACTCTTT	1134
Qy	784	GC 785	
Db	1135	GC 1136	

RESULT 15
AX794892
LOCUS AX794892 2433 bp DNA linear PAT 04-OCT-2003
DEFINITION Sequence 49 from Patent WO03051912.
ACCESSION AX794892
VERSION AX794892.1 GI:37515781
KEYWORDS Hepatitis C virus
SOURCE Hepatitis C virus
ORGANISM Hepatitis C virus
Virus; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
REFERENCE 1
Maertens,G., Depla,B. and Bosman,P.
AUTHORS Purified Hepatitis C virus envelope proteins for diagnostic and
TITLE therapeutic use
JOURNAL Patent: WO 03051912-A 49 26-JUN-2003;
INNOGENETICS N.V. (BE)

FEATURES	Location/Qualifiers
source	1. .2433
	/organism="Hepatitis C virus"
	/mol_type="unassigned DNA"
	/db_xref="taxon:11103"
CDS	1. .2430
	/note="unnamed protein product"
	/codon_start=1
	/protein_id="CAB48259.1"
	/db_xref="GI:37515782"
	/translation="MSTNPKPQRTKENTNERPODVKPPGGGQIVGGVLLPRRGPRLL GVRATTSRSRQPRGRQPIKAPRPBGRAWAOPGYPMPLVGNMGWAGWALLSPRG SRPSWPTDPRRRSNGKVIDTLCGFADLVGYIPLVAGPLAGGAALAHGVRVLED GVNATNGLPCGFSIFILLALLSCLTTPASAYEVNRVSGWVTHNDCNSISIVBAAD MIHTPCVPCVRENSSRCVWALTPTLAARNASVPTTIRRHVDLLVGAAPFCSAMT SOLLGCVFLVSOLFPTISPRHETVQDCNCSIYPGHTGRHMMNMNWNWSPPTALVV QDGLRIPQAVVUMVAGAHGVLGCLAYVSMVGNWAKVLVLMFAGVDGHTRVSGAA ASDTRGLVSLFSPGSAQKIQLVNTGWSHINRTALNCNDSLTQTFPALFPYKHENSS GCPRLASCSIDKFAQWGLPTTYTPNSSDRPFYCMHAPRPQCGIVPASQVCGPVYC FTSPVVGTTDRFQVPTYNWANDSVLLINLNRPPRNNWFGCTMNGTGTFTKTCGG PCNIGAGNNLTCTPTDCPRKHPEATVARGSGPMLTPRCMVHYPYRLWHVPTVNF TIPKRVYGVGVHRFEAACNWTGRBERCDLDRDRSELSPLLSTTTRWQIILPCSPPTLL PALSTGLIHLHONTVDVQYLYGVCSAVVSLVWKHEYVLLPLLADARIACACIWNMLL IAQRAALENLIVLNAARVAGAHGTLFLVFFCAANYIKGRLVPGAATAYFTGWNELL LLALAPPRAVA"
	1. .2427
	/product="unnamed"

mat_peptide	1. .2427
ORIGIN	
Query Match	98.0%; Score 778.8; DB 6; Length 2433;
Best Local Similarity	99.7%; Pred. No. 1.2e-163;
Matches	780; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy	4 TTGGTAAAGTTCATCGATACCTTTACATGCGGCTTCGCGACCTCGTGGGTACATTCG 63
Db	355 TTGGTAAAGTTCATCGATACCTTTACATGCGGCTTCGCGACCTCGTGGGTACATTCG 414
Qy	64 CTCGTGCGGCGCCCTTAGGGGGCGCTGCCAGGCGCTTGGCGCATGCGGTTCG 123
Db	415 CTCGTGCGGCGCCCTTAGGGGGCGCTGCCAGGCGCTTGGCGCATGCGGTTCG 474
Qy	124 GAGAGCGCGTGAACTATGCAACAGGGAATTTGCCCGGTTCCTTCTCTATCTTCTC 183
Db	475 GAGAGCGCGTGAACTATGCAACAGGGAATTTGCCCGGTTCCTTCTCTATCTTCTC 534
Qy	184 TTGGCTTTGCTGCTCTGCTGCAACCGTTCAGCTTCGCTTATGAGTGGCGAAGCTGTC 243
Db	535 TTGGCTTTGCTGCTCTGCTGCAACCGTTCAGCTTCGCTTATGAGTGGCGAAGCTGTC 594
Qy	244 GGGATGTACCATGTTCACGACGACTGCTCCAACTCAAGCATTTGTATGAGCGACGGAC 303
Db	595 GGGATGTACCATGTTCACGACGACTGCTCCAACTCAAGCATTTGTATGAGCGACGGAC 654
Qy	304 ATGATCATGACACCCCGGGTGGTGGCTTCCGCTTGGGAGAACAACTCTTCCGCTGC 363
Db	655 ATGATCATGACACCCCGGGTGGTGGCTTCCGCTTGGGAGAACAACTCTTCCGCTGC 714
Qy	364 TGGGTAGCGCTCACCCGCCACGCTGCGAGTGGCAAGCGCAGCGTCCGCCACACAAATA 423
Db	715 TGGGTAGCGCTCACCCGCCACGCTGCGAGTGGCAAGCGCAGCGTCCGCCACACAAATA 774
Qy	424 CGACGCCACGTCGATTTGCTGTTGGGGCGGCTGCTTTCTGCTGCTATGTACGTGGG 483
Db	775 CGACGCCACGTCGATTTGCTGTTGGGGCGGCTGCTTTCTGCTGCTATGTACGTGGG 834
Qy	484 GACCTCTGCGGATCTGCTTCTCGTCTCCAGCTGTTTCCACATCTCGCTCGCGCAT 543
Db	835 GACCTCTGCGGATCTGCTTCTCGTCTCCAGCTGTTTCCACATCTCGCTCGCGCAT 894
Qy	544 GAGCGGTGACGAGTGCATTTGCTCAATCTATCCCGGCCACATPACGGGTTCACCGTATG 603
Db	895 GAGCGGTGACGAGTGCATTTGCTCAATCTATCCCGGCCACATPACGGGTTCACCGTATG 954

QY 604 GCTTGGGATATGATGATGAACCTGGTCGCCCTACAAACGGCCCTGGTGGTATCGCAGCTGCTC 663
 Db |||||
 QY 955 GCTTGGGATATGATGATGAACCTGGTCGCCCTACAAACGGCCCTGGTGGTATCGCAGCTGCTC 1014
 Db |||||
 QY 664 CGGATCCCAACAGCTCTCGTGGACATGTTGGCGGGGCCCATTTGGGGAGTCTTGGCGGGT 723
 Db |||||
 QY 1015 CGGATCCCAACAGCTCTCGTGGACATGTTGGCGGGGCCCATTTGGGGAGTCTTGGCGGGC 1074
 Db |||||
 QY 724 CTCGCCCTACTATTCCATGGTGGGGAACCTGGGCTTAAGGTTTTCATTTGATGATGCTACTCTTT 783
 Db |||||
 QY 1075 CTCGCCCTACTATTCCATGGTGGGGAACCTGGGCTTAAGGTTTTCATTTGATGATGCTACTCTTT 1134
 Db |||||
 QY 784 GC 785
 Db ||
 QY 1135 GC 1136

Search completed: June 16, 2004, 11:39:16
 Job time : 2427.41 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2004, 07:39:53 ; Search time 299.119 Seconds
(without alignments)
11290.892 Million cell updates/sec

Title: US-09-899-303a-5

Perfect score: 795

Sequence: 1 ATGTTGGTAAGGTCATCGA.....TACTCTTGTGCTCCTAATAG 795

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_29Jan04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002s:*
- 7: Geneseqn2003as:*
- 8: Geneseqn2003bs:*
- 9: Geneseqn2003cs:*
- 10: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	795	100.0	795	2	AAT12705 HCV E1 CO
2	795	100.0	795	6	AAL48914 Hepatitis
3	795	100.0	795	9	ADD55513 Hepatitis
4	778.8	98.0	2082	6	AAL48939 Hepatitis
5	778.8	98.0	2082	9	ADD55555 Hepatitis
6	778.8	98.0	2086	2	AAT12973 HCV E1 CO
7	778.8	98.0	2433	2	AAT12974 HCV E1 CO
8	778.8	98.0	2433	9	ADD55557 Hepatitis
9	767.8	96.6	2434	6	AAL48940 Hepatitis
10	703.6	88.5	9605	6	ABK91431 Hepatitis
11	703.6	88.5	9605	6	ABK91424 Hepatitis
12	703.6	88.5	9605	6	ABK91429 Hepatitis
13	703.6	88.5	9605	6	ABK91432 Hepatitis
14	703.6	88.5	9605	6	ABK91411 Hepatitis
15	703.6	88.5	9605	6	ABK91430 Hepatitis
16	703.6	88.5	9605	6	ABK91428 Hepatitis
17	703.6	88.5	9605	6	ABK91425 Hepatitis
18	703.6	88.5	9605	6	ABK91426 Hepatitis
19	703.6	88.5	9605	6	ABK91433 Hepatitis
20	703.6	88.5	9605	6	AAD25332 Hepatitis
21	703.6	88.5	9608	6	ABK91427 Hepatitis
22	703.6	88.5	11062	6	AAD25331 Hepatitis
23	703.6	88.5	11076	3	AAA98965 Hepatitis

ALIGNMENTS

RESULT 1

AAT12705
ID AAT12705 standard; DNA; 795 BP.

XX AC AAT12705;

DT 23-SEP-1996 (first entry)

XX DB HCV E1 construct HCC110A.

KW HCV; E1; E2: disulphide bond cleavage; envelope protein; vaccine; human;
KW serotype; reversed phase hybridisation assay; genotype; antigen; sera;
KW ss.

XX OS Hepatitis C virus.

PN MO9604385-A2.

XX PD 15-FEB-1996.

XX PF 31-JUL-1995; 95WO-EP003031.

XX PR 29-JUL-1994; 94EP-00870132.

XX PA (INNO-) INNOGENETICS NV.

XX PI Maertens G, Bosman F, De Martynoff G, Buysse M;

XX DR WPI; 1996-129401/13.

XX PT Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope proteins
in presence of di-sulphide bond cleavage agent, to produce proteins
suitable for direct use in vaccines or diagnostic assays of HCV.

XX PS Claim 23; Fig 21; 146pp; English.

XX CC AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1
and E2 protein coding sequence constructs. These sequences are included
in vectors for the production of recombinant E1, E2, and E1/E2 proteins.
The recombinant proteins can then be isolated using a method of the
invention. In the method, the envelope proteins are purified by carrying
out a disulphide bond cleavage, or a reduction step with a disulphide
bond cleavage agent, after lysis of recombinant host cells. The
CC constructs containing the purified HCV envelope proteins can be used for
vaccinating humans against HCV, for in vitro detection of HCV antibodies
in a sample, and in a serotyping assay for detecting one or more

ABA03491 Cuticle p
Aa433889 NANB hepa
Aa633753 NANBHV ge
Aa24467 NANB hepa
Aa64068 Non-A, no
Aa30386 5'UTR/COR
Aa60672 Fragment
Aa155222 Plasmid p
Aa29628 Hepatitis
Aa60673 Fragment
Aa26981 HCV gene
Aa03677 Hepatitis
Aa80498 DNA encod
Aa81559 Hepatitis
Aa03960 Partial H
Aa25517 Hepatitis
Aa153723 Hepatitis
Aa49655 Hepatitis
Aa15363 Fragment
Aa12242 Encodes P
Aa20923 C10-E12 D
Aa33038 HCV-S1 fu

serological types of HCV present in a biological sample. The constructs can also be immobilised on a solid substrate and incorporated into a reversed phase hybridisation assay for determining the presence or the genotype of HCV. The new purification method preserves the conformation of the recombinantly expressed E1, E2 and E1/E2, and eliminates contaminating proteins. Antigens isolated using this method are more reactive with human sera than those isolated by known techniques

Sequence 795 BP: 130 A: 240 C: 231 G: 194 T: 0 U: 0 Other: 0

```
Query Match      100.0%; Score 795; DB 2; Length 795;
Best Local Similarity 100.0%; Pred. No. 4.1e-204;
Matches 795; Conservative 0; Mismatches 0; Indels
```

Qy	1	ATGTTGGGTAAAGTTCATGATACCTTTACATGCGGCTTCGCGGACCTCGTGGGGTACATT	60
Db	1	ATGTTGGGTAAAGTTCATGATACCTTTACATGCGGCTTCGCGGACCTCGTGGGGTACATT	60
Qy	61	CCGCTCGTGGCGGCCCCCTTAGGGGGCGCTGCAGGGGCCCTGGCGCATGGGCTCCGGGTT	120
Db	61	CCGCTCGTGGCGGCCCCCTTAGGGGGCGCTGCAGGGGCCCTGGCGCATGGGCTCCGGGTT	120
Qy	121	CTGAGGACGGCGTGAACCTATGCAACAGGGAATTTGCCCGGTGCTCTTTCTATCTTC	180
Db	121	CTGAGGACGGCGTGAACCTATGCAACAGGGAATTTGCCCGGTGCTCTTTCTATCTTC	180
Qy	181	CTCTTGGCTTTGCTGTCTGTGACCGTTCAGCTTCGGCTTATGAAGTGGCGAACGTG	240
Db	181	CTCTTGGCTTTGCTGTCTGTGACCGTTCAGCTTCGGCTTATGAAGTGGCGAACGTG	240
Qy	241	TCGGGATGTACCATGTTCACGAACGACTGCTCAACTCAAGCATTTGCTATGAGGACGCG	300
Db	241	TCGGGATGTACCATGTTCACGAACGACTGCTCAACTCAAGCATTTGCTATGAGGACGCG	300
Qy	301	GACATGATCATGCAACCCCCGGGTGGTGCCTGCTTGGGAGAAACACTTTTCCGCG	360
Db	301	GACATGATCATGCAACCCCCGGGTGGTGCCTGCTTGGGAGAAACACTTTTCCGCG	360
Qy	361	TGCTGGGTAGGCTCACCCCAACGCTCGACGTAGGAACGCCAGGTCGCCACACGACA	420
Db	361	TGCTGGGTAGGCTCACCCCAACGCTCGACGTAGGAACGCCAGGTCGCCACACGACA	420
Qy	421	ATACGACGCCACGTGCGATTTGCTGCTTGGGGCGGCTGCTTTCTGTTCCGCTATGTACGTG	480
Db	421	ATACGACGCCACGTGCGATTTGCTGCTTGGGGCGGCTGCTTTCTGTTCCGCTATGTACGTG	480
Qy	481	GGGGAOCTCTGGCGGATCTGTCTTCTCTCTCCAGCTGTTTCAACATCTCGCCTCGCCGG	540
Db	481	GGGGAOCTCTGGCGGATCTGTCTTCTCTCTCCAGCTGTTTCAACATCTCGCCTCGCCGG	540
Qy	541	CATGAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCGGCCACATAACGGGTCAACGT	600
Db	541	CATGAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCGGCCACATAACGGGTCAACGT	600
Qy	601	ATGCTTTGGGATATGATGAACTGTGCGCTTCAACGCGGCCCTGGTGGTATGCGAGCTG	660
Db	601	ATGCTTTGGGATATGATGAACTGTGCGCTTCAACGCGGCCCTGGTGGTATGCGAGCTG	660
Qy	661	CTCCGGATCCACAAGCTGTCTGTGACATGTGTGGCGGGGCCCATTTGGGGAGTCTGGCG	720
Db	661	CTCCGGATCCACAAGCTGTCTGTGACATGTGTGGCGGGGCCCATTTGGGGAGTCTGGCG	720
Qy	721	GGTCTCGCTACTATTCCATGTTGGGGAACCTGGGCTTAAGGTTTTCGATGTGATCTACTTC	780
Db	721	GGTCTCGCTACTATTCCATGTTGGGGAACCTGGGCTTAAGGTTTTCGATGTGATCTACTTC	780
Qy	781	TTTGCTCCCTAATAG	795
Db	781	TTTGCTCCCTAATAG	795

RESULT 2
AAL48914

ID	AA148914	standard; DNA; 795 BP.
XX	AC	
XX	AC	AA148914;
XX	XX	
DT	24-OCT-2002	(first entry)
DT	XX	
XX	XX	Hepatitis C virus clone HCC110A E1 protein coding sequence.
DE	DE	
XX	XX	
KW	KW	Hepatitis C virus; HCV; E1 protein; E2 protein; infection; gene;
KW	KW	virucide; immunostimulant; vaccine; ds.
XX	XX	
OS	OS	Hepatitis C virus.
XX	XX	
PN	PN	WO200255548-A2.
XX	XX	
XX	XX	18-JUL-2002.
PD	XX	
XX	XX	
PF	PF	11-JAN-2002; 2002WO-EP000219.
XX	XX	
XX	XX	
PR	PR	11-JAN-2001; 2001US-0260669P.
PR	PR	30-AUG-2001; 2001US-0315768P.

```
Query Match      100.0%; Score 795; DB 6; Length 795;
Best Local Similarity 100.0%; Pred. No. 4.1e-204;
Matches 795; Conservative 0; Mismatches 0; Indels 0; Gaps 0
```

Qy	1	ATGTCGGTAAAGTCAATCGATACCCCTTACATGCGCGCTTCGCCGACCTCGTGGGGTACATT	60
Db	1	ATGTCGGTAAAGTCAATCGATACCCCTTACATGCGCGCTTCGCCGACCTCGTGGGGTACATT	60
Qy	61	CCGCTCGTCGGCGCCCCCTTAGGGGGCGCTGCCACAGGCGCCCTGGCGGCATGGCGTCCGGGTT	120
Db	61	CCGCTCGTCGGCGCCCCCTTAGGGGGCGCTGCCACAGGCGCCCTGGCGGCATGGCGTCCGGGTT	120
Qy	121	CTGGGGAACGGCGTAGCAATATGCAACAGGGAATTTGCCCGGTCTCTTTCTCTATATCTTC	180
Db	121	CTGGGGAACGGCGTAGCAATATGCAACAGGGAATTTGCCCGGTCTCTTTCTCTATATCTTC	180
Qy	181	CTCTTGCGTTTGCTGTCTGTCTGACCGGTTCCAGCTTCGGCTTATGAAGTGGCGAACGTG	240
Db	181	CTCTTGCGTTTGCTGTCTGTCTGACCGGTTCCAGCTTCGGCTTATGAAGTGGCGAACGTG	240
Qy	241	TCCGGGATGTACATGTCAAGCAAGACTGCTCCAACCTCAAGCATTTGTGTATGAGCGCAGCG	300

```
Db 241 TCCGGATGATACATGTCAAGACGACTGCTCCAACTCAAGCATTTGTGTATGAGCAGG 300
Qy 301 GACATGATCATGACACCCCGGGTGGTGGCTGCGCTTGGGAGAAACAATCTTTCCCGC 360
Db 301 GACATGATCATGACACCCCGGGTGGTGGCTGCGCTTGGGAGAAACAATCTTTCCCGC 360
Qy 361 TGCTGGGTAGCGCTCACCCACACGCTCCGAGCTAGGAAAGCGCAGCGTCCCAACACGACA 420
Db 361 TGCTGGGTAGCGCTCACCCACACGCTCCGAGCTAGGAAAGCGCAGCGTCCCAACACGACA 420
Qy 421 ATACGAGCCACGCTGATTTGCTCGTTGGGGGGCTGCTTTCTGTTCGGCTATGTACGTG 480
Db 421 ATACGAGCCACGCTGATTTGCTCGTTGGGGGGCTGCTTTCTGTTCGGCTATGTACGTG 480
Qy 481 GGGGACCTCTCGGATCTGCTCTCTCTCCAGCTGTTCCAGCTGTTCCACATCTCGCTCGCGG 540
Db 481 GGGGACCTCTCGGATCTGCTCTCTCTCCAGCTGTTCCACATCTCGCTCGCGG 540
Qy 541 CATGAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTCAACGT 600
Db 541 CATGAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTCAACGT 600
Qy 601 ATGGCTTTGGGATATGATGATGAATGCTGCTGCTTACCAAGCGCTGCTGCTGATCGCAGCTG 660
Db 601 ATGGCTTTGGGATATGATGATGAATGCTGCTGCTTACCAAGCGCTGCTGCTGATCGCAGCTG 660
Qy 661 CTCCGGATCCCAAGCTGCTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
Db 661 CTCCGGATCCCAAGCTGCTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
Qy 721 GGTCTCGCTTACTATTCCATGCTGGGAACTGGGCTAAGGTTTGTATGATGCTACTC 780
Db 721 GGTCTCGCTTACTATTCCATGCTGGGAACTGGGCTAAGGTTTGTATGATGCTACTC 780
Qy 781 TTTGCTCCCTAATAG 795
Db 781 TTTGCTCCCTAATAG 795
```

RESULT 3

```
ADD55513
ID ADD55513 standard; DNA; 795 BP.
XX
AC ADD55513;
XX
DT 15-JAN-2004 (first entry)
XX
DE Hepatitis C virus E1 protein coding sequence #2.
XX
KW Hepatitis C virus; HCV; vaccine; liver disease; E1 protein; E2 protein;
XX liver fibrosis; ds; gene.
XX
OS Hepatitis C virus.
XX
PN WO2003051912-A2.
XX
PD 26-JUN-2003.
XX
PF 18-DEC-2002; 2002WO-EP014480.
XX
PR 18-DEC-2001; 2001US-00020510.
XX
PR 16-OCT-2002; 2002US-0418358P.
XX
XX (INNO-) INNOGENETICS NV.
XX
XX Maertens G, Depla E, Bosman F;
XX
XX WPI; 2003-541632/51.
XX
XX P-PSDB; ADD55514.
XX
XX New hepatitis C virus (HCV) vaccine composition, useful for reducing
XX liver disease, e.g., liver fibrosis in a chronic HCV-infected mammal.
```

PS Example 1; SEQ ID NO 5; 271pp; English.

XX The invention comprises an Hepatitis C virus (HCV) vaccine for reducing liver disease. The vaccine of the invention comprises an HCV E1 or E2 protein as an antigen. The HCV vaccine is useful for reducing liver disease (e.g. liver fibrosis) in a chronic HCV-infected mammal. The present DNA sequence encodes an HCV E1 protein.

XX Sequence 795 BP; 130 A; 240 C; 231 G; 194 T; 0 U; 0 Other;

```
Query Match 100.0%; Score 795; DB 9; Length 795;
Best Local Similarity 100.0%; Pred. No. 4.1e-204;
Matches 795; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGTTGGTAAAGTTCATCATGATACCTTATCATGCGCTTTCGCGAGCTCTGTGGGGTACATT 60
Db 1 ATGTTGGTAAAGTTCATCATGATACCTTATCATGCGCTTTCGCGAGCTCTGTGGGGTACATT 60
Qy 61 CCGCTCGTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 120
Db 61 CCGCTCGTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 120
Qy 121 CTGGAGGACGCGGTGAACATATGCAACAGGAAATTTGCCGGTTCCTTTCTCTATCTTC 180
Db 121 CTGGAGGACGCGGTGAACATATGCAACAGGAAATTTGCCGGTTCCTTTCTCTATCTTC 180
Qy 181 CTCTTGGCTTTTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
Db 181 CTCTTGGCTTTTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
Qy 241 TCCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 300
Db 241 TCCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 300
Qy 301 GACATGATCATGACACCCCGGGTGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
Db 301 GACATGATCATGACACCCCGGGTGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
Qy 361 TGCTGGGTAGCGCTCACCCACACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
Db 361 TGCTGGGTAGCGCTCACCCACACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
Qy 421 ATACGAGCCACGCTGATTTGCTCGTTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 480
Db 421 ATACGAGCCACGCTGATTTGCTCGTTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 480
Qy 481 GGGGACCTCTCGGATCTGCTCTCTCTCCAGCTGTTCCAGCTGTTCCACATCTCGCTCGCGG 540
Db 481 GGGGACCTCTCGGATCTGCTCTCTCTCCAGCTGTTCCACATCTCGCTCGCGG 540
Qy 541 CATGAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTCAACGT 600
Db 541 CATGAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTCAACGT 600
Qy 601 ATGGCTTTGGGATATGATGATGAATGCTGCTGCTTACCAAGCGCTGCTGCTGATCGCAGCTG 660
Db 601 ATGGCTTTGGGATATGATGATGAATGCTGCTGCTTACCAAGCGCTGCTGCTGATCGCAGCTG 660
Qy 661 CTCCGGATCCCAAGCTGCTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
Db 661 CTCCGGATCCCAAGCTGCTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
Qy 721 GGTCTCGCTTACTATTCCATGCTGGGAACTGGGCTAAGGTTTGTATGATGCTACTC 780
Db 721 GGTCTCGCTTACTATTCCATGCTGGGAACTGGGCTAAGGTTTGTATGATGCTACTC 780
Qy 781 TTTGCTCCCTAATAG 795
Db 781 TTTGCTCCCTAATAG 795
```

RESULT 4

AAL48939

ID AAL48939 standard; DNA; 2082 BP.
 AC AAL48939;
 DT 24-OCT-2002 (first entry)
 DE Hepatitis C virus E2 protein related coding sequence SEQ ID NO: 47.
 KW Hepatitis C virus; HCV; E1 protein; E2 protein; infection; gene;
 KW virucide; immunostimulant; vaccine; ds.
 XX
 OS Hepatitis C virus.
 XX
 PN WO20025548-A2.
 XX
 PD 18-JUL-2002.
 XX
 PF 11-JAN-2002; 2002WO-EP000219.
 XX
 PR 11-JAN-2001; 2001US-0260669P.
 PR 30-AUG-2001; 2001US-0315768P.
 XX
 XX (INNO-) INNOGENETICS NV.
 PA
 XX Maertens G, Bosman F, Buyse M;
 PI P-PSDB; AAO18678.
 DR
 DR WPI; 2002-599657/64.
 XX P-PSDB; AAO18678.
 XX
 PT New therapeutic vaccine compositions comprising at least one purified
 PT recombinant hepatitis C virus (HCV) single or specific oligomeric
 PT recombinant envelope protein E1 or E2, useful for immunizing humans from
 PT HCV infection.
 XX
 PS Example 2; Page 206-209; 243pp; English.
 XX
 CC The present invention relates to new therapeutic vaccine compositions for
 CC inducing Hepatitis C virus (HCV)-specific antibodies, comprising a
 CC composition containing at least one purified recombinant HCV single or
 CC specific oligomeric recombinant envelope proteins selected from an E1 and
 CC an E2 protein, and optionally a pharmaceutical adjuvant. The vaccines are
 CC useful for inducing HCV-specific antibodies or for immunising humans
 CC against HCV. The recombinant HCV E1 and/or E2 proteins are useful as
 CC vaccines or therapeutics, in HCV screening and confirmatory antibody
 CC tests, for raising antibodies, in the preparation of medicament, and for
 CC in vitro monitoring of HCV disease or prognosis of the response to
 CC treatment of patients suffering from HCV infection. The present sequence
 CC is a coding sequence described in the exemplification of the invention
 XX
 SQ Sequence 2082 BP; 366 A; 634 C; 600 G; 482 T; 0 U; 0 Other;
 Query Match 98.0%; Score 778.8; DB 6; Length 2082;
 Best Local Similarity 99.7%; Pred. No. 1.3e-199;
 Matches 780; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 OY 4 TTGGGTAAAGTTCATGATACCTTACATCGCGCTTCGCGACCTCGTGGGTACATTCG 63
 DB 4 TTGGGTAAAGTTCATGATACCTTACATCGCGCTTCGCGACCTCGTGGGTACATTCG 63
 OY 64 CTCGTGCGCGCCCCCTAGGGGGCGCTGCGAGGGCCCTGCGCGATGCGGTTCG 123
 DB 64 CTCGTGCGCGCCCCCTAGGGGGCGCTGCGAGGGCCCTGCGCGATGCGGTTCG 123
 OY 124 GAGGACGGGTGATATGATGACACAGGAAATTTCCCGGTGCTTCTTCTATCTTCCTC 183
 DB 124 GAGGACGGGTGATATGATGACACAGGAAATTTCCCGGTGCTTCTTCTATCTTCCTC 183
 OY 184 TTGGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 243
 DB 184 TTGGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 243
 OY 244 GGGATGTATCATGTCAAGCACTGCTCCAACTCAAGCAATGTGTATGAGGCGGCGAC 303
 DB 244 GGGATGTATCATGTCAAGCACTGCTCCAACTCAAGCAATGTGTATGAGGCGGCGAC 303

Db 244 GGGATGTATCATGTCAAGCACTGCTCCAACTCAAGCAATGTGTATGAGGCGGCGAC 303
 OY 304 ATGATCATGTCAAGCACTGCTCCAACTCAAGCAATGTGTATGAGGCGGCGAC 363
 Db 304 ATGATCATGTCAAGCACTGCTCCAACTCAAGCAATGTGTATGAGGCGGCGAC 363
 OY 364 TTGGGTAGCGCTCAAGCACTGCTCCAACTCAAGCAATGTGTATGAGGCGGCGAC 423
 Db 364 TTGGGTAGCGCTCAAGCACTGCTCCAACTCAAGCAATGTGTATGAGGCGGCGAC 423
 OY 424 CGAGCCGACGTGATTTGCTCGTTGGGGGGGGTCTTCTGTTCCGCTATGTAAGTGGG 483
 Db 424 CGAGCCGACGTGATTTGCTCGTTGGGGGGGGTCTTCTGTTCCGCTATGTAAGTGGG 483
 OY 484 GACCTCTGCGGATCTGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 543
 Db 484 GACCTCTGCGGATCTGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 543
 OY 544 GAGAGGTGCGAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTCAACGGTATG 603
 Db 544 GAGAGGTGCGAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAACGGGTCAACGGTATG 603
 OY 604 GCTTGGGATATGATGATGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 663
 Db 604 GCTTGGGATATGATGATGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 663
 OY 664 CGGATCCCAACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 723
 Db 664 CGGATCCCAACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 723
 OY 724 CTCGCTACTATTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 783
 Db 724 CTCGCTACTATTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 783
 OY 784 GC 785
 Db 784 GC 785
 RESULT 5
 ID ADD55555 standard; DNA; 2082 BP.
 AC ADD55555;
 XX
 DT 15-JAN-2004 (first entry)
 XX
 DE Hepatitis C virus E1/E2 protein coding sequence #13.
 KW Hepatitis C virus; HCV; vaccine; liver disease; E1 protein; E2 protein;
 KW liver fibrosis; ds; gene.
 XX
 OS Hepatitis C virus.
 XX
 PN WO2003051912-A2.
 XX
 PD 26-JUN-2003.
 XX
 PF 18-DEC-2002; 2002WO-EP014480.
 XX
 PR 18-DEC-2001; 2001US-00020510.
 PR 16-OCT-2002; 2002US-0418358P.
 XX
 XX (INNO-) INNOGENETICS NV.
 PA
 XX Maertens G, Depla B, Bosman F;
 PI P-PSDB; ADD55555.
 DR
 DR WPI; 2003-541632/S1.
 XX P-PSDB; ADD55555.
 XX
 PT New hepatitis C virus (HCV) vaccine composition, useful for reducing
 PT liver disease, e.g., liver fibrosis in a chronic HCV-infected mammal.
 XX

[illegible]

AT12973	standard; DNA; 2086 BP.
AT12973;	
24-SEP-1996	(first entry)
HCV E1 construct	HCI165.
HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human; serotype; reversed phase hybridisation assay; genotype; antigen; sera; ss.	
Hepatitis C virus.	
WD9604385-A2.	
15-FEB-1996.	
31-JUL-1995;	95WO-EP003031.
29-JUL-1994;	94EP-00870132.
(INNO-) INNOGENETICS NV.	
Maertens G, Bosman P, De Martynoff G, Buyse M;	
WPI; 1996-129401/13.	
Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope proteins - in presence of di:sulphide bond cleavage agent, to produce proteins suitable for direct use in vaccines or diagnostic assays of HCV.	
Claim 23; Fig 21; 146pp; English.	
AT12704-T12709 and AT12961-T12974 represent hepatitis C virus (HCV) E1 and E2 protein coding sequence constructs. These sequences are included in vectors for the production of recombinant E1, E2, and E1/E2 proteins. The recombinant proteins can then be isolated using a method of the invention. In the method, the envelope proteins are purified by carrying out a disulphide bond cleavage, or a reduction step with a disulphide bond cleavage agent, after lysis of recombinant host cells. The constructs containing the purified HCV envelope proteins can be used for vaccinating humans against HCV, for in vitro detection of HCV antibodies in a sample, and in a serotyping assay for detecting one or more serological types of HCV present in a biological sample. The constructs can also be immobilised on a solid substrate and incorporated into a reversed phase hybridisation assay for determining the presence or the genotype of HCV. The new purification method preserves the conformation of the recombinantly expressed E1, E2 and E1/E2, and eliminates contaminating proteins. Antigens isolated using this method are more reactive with human sera than those isolated by known techniques	
Sequence 2086 BP; 366 A; 635 C; 601 G; 484 T; 0 U; 0 Other;	

QY 664 CGATCCCAAGCTGCTGTCGACATGTCGCGGGGCCCAATTGGGGAGTCTCTGGCGGGT 723
 DB 1015 CGGATCCCAAGCTGCTGTCGACATGTCGCGGGGCCCAATTGGGGAGTCTCTGGCGGGC 1074

QY 724 CTGCGCTACTATTCATGCTGGGAACTGGGCTAAGGCTTTTGGATGTGATGCTACTCTTT 783
 DB 1075 CTGCGCTACTATTCATGCTGGGAACTGGGCTAAGGCTTTTGGATGTGATGCTACTCTTT 1134

QY 784 GC 785
 DB 1135 GC 1136

RESULT 8
 ADD55557
 ID ADD55557 standard; DNA; 2433 BP.
 AC ADD55557;
 XX
 DT 15-JAN-2004 (first entry)
 XX
 DE Hepatitis C virus E1/E2 protein coding sequence #14.
 XX
 KW Hepatitis C virus; HCV; vaccine; liver disease; E1 protein; E2 protein;
 KW liver fibrosis; ds; gene.
 XX
 OS Hepatitis C virus.
 XX
 PN WO2003051912-A2.
 XX
 PD 26-JUN-2003.
 XX
 PP 18-DEC-2002; 2002WO-BP014480.
 XX
 PR 18-DEC-2001; 2001US-00020510.
 PR 16-OCT-2002; 2002US-0418358P.
 XX
 PA (INNO-) INNOGENETICS NV.
 XX
 PI Maertens G, Depla E, Bosman F;
 XX
 DR WPI; 2003-541632/51.
 DR P-PSDB; ADD55558.
 XX
 PT New hepatitis C virus (HCV) vaccine composition, useful for reducing
 PT liver disease, e.g., liver fibrosis in a chronic HCV-infected mammal.
 XX
 PS Example 2; SEQ ID NO 49; 271pp; English.
 XX
 CC The invention comprises an Hepatitis C virus (HCV) vaccine for reducing
 CC liver disease. The vaccine of the invention comprises an HCV E1 or E2
 CC protein as an antigen. The HCV vaccine is useful for reducing liver
 CC disease (e.g. liver fibrosis) in a chronic HCV-infected mammal. The
 CC present DNA sequence encodes an HCV E1/E2 protein.
 XX
 SQ Sequence 2433 BP; 434 A; 745 C; 714 G; 540 T; 0 U; 0 Other;
 Query Match 98.0%; Score 778.8; DB 9; Length 2433;
 Best Local Similarity 99.7%; Pred. No. 1.4e-199;
 Matches 780; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 TTGGGTAAGGTATCGATACCTTACATGCGGCTTCGCGACCTCGTGGGGTACATTCOG 63
 DB 355 TTGGGTAAGGTATCGATACCTTACATGCGGCTTCGCGACCTCGTGGGGTACATTCOG 414

QY 64 CTGCTCGGCGCCCTAGGCGGCGCTGCCAGGCGCTGGCGCATGGCGTCCGGGTTCTG 123
 DB 415 CTGCTCGGCGCCCTAGGCGGCGCTGCCAGGCGCTGGCGCATGGCGTCCGGGTTCTG 474

QY 124 GAGGACGGCGTGAATATGCAACAGGGAATTTGCGCGGTTGCTCTTTCTATCTCTC 183
 DB 475 GAGGACGGCGTGAATATGCAACAGGGAATTTGCGCGGTTGCTCTTTCTATCTCTC 534

QY 184 TTGGCTTTGCTGCTGCTGCTGACCGTTCAGCTTCGGCTTATGAAGTGGCAACGTGTC 243
 DB 535 TTGGCTTTGCTGCTGCTGCTGACCGTTCAGCTTCGGCTTATGAAGTGGCAACGTGTC 594

QY 244 GGGATGTAACCATGTCACGACGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 303
 DB 595 GGGATGTAACCATGTCACGACGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 654

QY 304 ATGATCATGTCACACCCCGGGTGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 363
 DB 555 ATGATCATGTCACACCCCGGGTGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 714

QY 364 TGGGTAGCGCTCACCCCGGGTGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 423
 DB 715 TGGGTAGCGCTCACCCCGGGTGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 774

QY 424 CGACGCCACGTCGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 483
 DB 775 CGACGCCACGTCGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 834

QY 484 GACCTCTGCGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 543
 DB 835 GACCTCTGCGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 894

QY 544 GAGACGGTGCAGGACTGCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 603
 DB 895 GAGACGGTGCAGGACTGCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 954

QY 604 GCTTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 663
 DB 955 GCTTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1014

QY 664 CGGATCCCAAGCTGCTGTCGACATGTCGCGGGGCCCAATTGGGGAGTCTCTGGCGGGT 723
 DB 1015 CGGATCCCAAGCTGCTGTCGACATGTCGCGGGGCCCAATTGGGGAGTCTCTGGCGGGC 1074

QY 724 CTGCGCTACTATTCATGCTGGGAACTGGGCTTAAGGCTTTTGGATGTGATGCTACTCTTT 783
 DB 1075 CTGCGCTACTATTCATGCTGGGAACTGGGCTTAAGGCTTTTGGATGTGATGCTACTCTTT 1134

QY 784 GC 785
 DB 1135 GC 1136

RESULT 9
 AAL48940
 ID AAL48940 standard; DNA; 2434 BP.
 AC AAL48940;
 XX
 DT 24-OCT-2002 (first entry)
 XX
 DE Hepatitis C virus E2 protein related coding sequence SEQ ID NO: 49.
 XX
 KW Hepatitis C virus; HCV; E1 protein; E2 protein; infection; gene;
 KW virucide; immunostimulant; vaccine; ds.
 XX
 OS Hepatitis C virus.
 XX
 PN WO200255548-A2.
 XX
 PD 18-JUL-2002.
 XX
 PP 11-JAN-2002; 2002WO-BP000219.
 XX
 PR 11-JAN-2001; 2001US-0260669P.
 PR 30-AUG-2001; 2001US-0315788P.
 XX
 PA (INNO-) INNOGENETICS NV.
 XX
 PI Maertens G, Bosman F, Buyse M;
 XX

DR WPI; 2002-599657/64.
 DR P-PSDB; AAO18679.
 XX New therapeutic vaccine compositions comprising at least one purified
 PT recombinant hepatitis C virus (HCV) single or specific oligomeric
 FT recombinant envelope protein E1 or E2, useful for immunizing from
 PT HCV infection.
 XX
 PS Example 2; Page 212-215; 243pp; English.
 XX
 CC The present invention relates to new therapeutic vaccine compositions for
 CC inducing hepatitis C virus (HCV)-specific antibodies, comprising a
 CC composition containing at least one purified recombinant HCV single or
 CC specific oligomeric recombinant envelope proteins selected from an E1 and
 CC an E2 protein, and optionally a pharmaceutical adjuvant. The vaccines are
 CC useful for inducing HCV-specific antibodies or for immunising humans
 CC against HCV. The recombinant HCV E1 and/or E2 proteins are useful as
 CC vaccines or therapeutics, in HCV screening and confirmatory antibody
 CC tests, for raising antibodies, in the preparation of medicament, and for
 CC in vitro monitoring of HCV disease or prognosing the response to
 CC treatment of patients suffering from HCV infection. The present sequence
 CC is a coding sequence described in the exemplification of the invention
 XX
 SQ Sequence 2434 BP; 434 A; 745 C; 714 G; 541 T; 0 U; 0 Other;
 Query Match 96.6%; Score 767.8; DB 6; Length 2434;
 Best Local Similarity 99.6%; Pred. No. 1.3e-196;
 Matches 780; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
 4 TTGGTAAGGTCATCGATACCTTATACATCGCGCTTGGCGACCTCGTGGGGTACATTCGG 63
 Db TTGGTAAGGTCATCGATACCTTATACATCGCGCTTGGCGACCTCGTGGGGTACATTCGG 414
 64 CTGTCGCGGCGCCCTAGGCGCGCTGCGAGGCGCTGCGGCGAGCGTCCGGGTTGCG 123
 Db CTGTCGCGGCGCCCTAGGCGCGCTGCGAGGCGCTGCGGCGAGCGTCCGGGTTGCG 474
 124 GAGGACGGGTGAATATGCAACAGGGAATTTGCGCGGTTGCTTTCTCTATCTTCCTC 183
 Db GAGGACGGGTGAATATGCAACAGGGAATTTGCGCGGTTGCTTTCTCTATCTTCCTC 534
 184 TTGGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 242
 Db TTGGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 594
 243 CGGATGTACCATGTGTCAGAAAGACGTCTCAACTCAAGCATTTGTGTATGAGCGAGCG 302
 Db CGGATGTACCATGTGTCAGAAAGACGTCTCAACTCAAGCATTTGTGTATGAGCGAGCG 654
 303 CATGATCATGACACCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 362
 Db CATGATCATGACACCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 714
 363 CTGGTAGCGCTCACCCACACCTGCGAGCTAGGAAAGCGGTCGCCACCAACGACAT 422
 Db CTGGTAGCGCTCACCCACACCTGCGAGCTAGGAAAGCGGTCGCCACCAACGACAT 774
 423 ACAGCGCCACGTCGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 482
 Db ACAGCGCCACGTCGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 834
 483 GGACCTCTGCGGATCTGCT 542
 Db GGACCTCTGCGGATCTGCT 894
 543 TGAGACGCTGCGAGCTGCAATGCTCAATCTATATCCCGGCCACATACCGGGTCACGTA 602
 Db TGAGACGCTGCGAGCTGCAATGCTCAATCTATATCCCGGCCACATACCGGGTCACGTA 954
 603 GCGCTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 662
 Db GCGCTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1014

QY 663 CCGGATCCCAACAGCTGTCTGTGACATGCTGGCGGGGCCCATTTGGGAGTCTCTGCGGG 722
 Db CCGGATCCCAACAGCTGTCTGTGACATGCTGGCGGGGCCCATTTGGGAGTCTCTGCGGG 1074
 QY 723 TCTCGCTTACTATTCATGCTGGGAACTGGGCTAAGGTTTTCATTTGATGCTACTCTT 782
 Db CTTGCTTACTATTCATGCTGGGAACTGGGCTAAGGTTTTCATTTGATGCTACTCTT 1134
 QY 783 TGC 785
 Db 1135 TGC 1137
 RESULT 10
 ID ABK91431 standard; DNA; 9605 BP.
 AC ABK91431;
 XX 15-NOV-2002 (first entry)
 XX Hepatitis C virus Con 1 isolate DNA mutant 8.
 XX HCV; ds; Con 1; adaptive mutation; liver failure; cirrhosis;
 KW hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;
 KW internal ribosome entry site; IRES; NS5A; HCV replication; mutant.
 OS Hepatitis C virus.
 OS Synthetic.
 XX Key Location/Qualifiers
 CDS 342..9374
 FT /*tag= a
 FT /product= "HCV polyprotein"
 FT /note= "The polyprotein consists of the Core, E1, E2, P7,
 FT NS2, NS3, NS4A, NS4B, NS5A and NS5B proteins"
 FT mutation replace(6934,T)
 FT /*tag= b
 WO200259321-A2.
 PD 01-AUG-2002.
 XX 16-JAN-2002; 2002MO-BP000526.
 XX 23-JAN-2001; 2001US-0263479P.
 XX (RICE-) IST RICERHE BIOL MOLECOLARE ANGELETTI.
 XX De Francesco R, Migliaccio G, Paonessa G;
 WPI; 2002-599793/64.
 XX New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV
 PT NS5 encoding region, or encephalomyocarditis virus (EMCV) internal
 PT ribosome entry site (IRES) region, useful in studying HCV replication and
 PT expression.
 XX Claim 9; Page; 69pp; English.
 PS
 XX The invention relates to nucleic acid molecules comprising altered HCV
 CC NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)
 CC internal ribosome entry site (IRES) region coding for one or more NS3,
 CC NS5A, or EMCV IRES mutations, respectively. The location of the mutations
 CC are detailed in the specification. Also included are (1) an expression
 CC vector comprising a nucleotide sequence coding for the altered nucleic
 CC acids, which is transcriptionally coupled to an exogenous promoter; (2) a
 CC recombinant cell human hepatoma cell comprising the altered nucleic acids
 CC ; (3) a recombinant cell produced by introducing an HCV (hepatitis C virus)
 CC cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)
 CC replicon enhanced cell or which containing a functional HCV replicon; (5)
 CC an HCV replicon enhanced cells made in the method; and (6) measuring the
 CC ability of a compound to affect HCV activity. The HCV replicons and HCV

CC replicon enhanced cells are useful in studying HCV replication and
 CC expression, and HCV and host cell interactions, producing HCV RNA and
 CC proteins, and providing a system for measuring the ability of a compound
 CC to modulate one or more HCV activities e.g. to discover drugs which may
 CC treat HCV mediated diseases such as liver failure, cirrhosis and
 CC hepatocellular carcinoma. The present sequence is an HCV replicon Con 1
 CC mutant of the invention. Note: The present sequence is not shown in the
 CC specification but was created by the indexer using the HCV sequence
 CC appearing as ABK91411 and the information in Claim 9
 XX
 SQ Sequence 9605 BP; 1910 A; 2884 C; 2733 G; 2078 T; 0 U; 0 Other;
 Query Match 88.5%; Score 703.6; DB 6; Length 9605;
 Best Local Similarity 93.7%; Pred. No. 4e-179;
 Matches 733; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
 QY 4 TTGGGTAAGTCAATGATACCTTACATGCGCTTTCGCCGACCTCGTGGGTACATCCG 63
 DB |||||
 QY 696 TTGGGTAAGTCAATGATACCTTACATGCGCTTTCGCCGACCTCGTGGGTACATCCG 755
 DB |||||
 QY 64 CTCGTGCGGCCCCCTAGGGGGCGCTGCCAGGGCCCTTGGCGCATGGCGTCCGGGTTCTG 123
 DB |||||
 QY 756 CTCGTGCGGCCCCCTAGGGGGCGCTGCCAGGGCCCTTGGCGCATGGCGTCCGGGTTCTG 815
 DB |||||
 QY 124 GAGGACGGGTGAATGATGCAAGGAAATTTGGCCGGTTGCTCTTCTATCTTCTCTC 183
 DB |||||
 QY 816 GAGGACGGGTGAATGATGCAAGGAAATTTGGCCGGTTGCTCTTCTATCTTCTCTC 875
 DB |||||
 QY 184 TTGGGTAAGTCAATGATACCTTACATGCGCTTTCGCCGACCTCGTGGGTACATCCG 243
 DB |||||
 QY 876 TTGGGTAAGTCAATGATACCTTACATGCGCTTTCGCCGACCTCGTGGGTACATCCG 935
 DB |||||
 QY 244 GGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 303
 DB |||||
 QY 936 GGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 995
 DB |||||
 QY 304 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 363
 DB |||||
 QY 996 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1055
 DB |||||
 QY 364 TGGGTAGCGCTCAACCCCGCGGTGCGTCTTCCAGCTTCCAGCTTCCAGCTTCCAGCTT 423
 DB |||||
 QY 1056 TGGGTAGCGCTCAACCCCGCGGTGCGTCTTCCAGCTTCCAGCTTCCAGCTTCCAGCTT 1115
 DB |||||
 QY 424 CGAGCGGACCGTCCGATTTGCTCGTGGGGCGGCTCTTCTTCTTCTTCTTCTTCTTCTT 483
 DB |||||
 QY 1116 CGAGCGGACCGTCCGATTTGCTCGTGGGGCGGCTCTTCTTCTTCTTCTTCTTCTTCTT 1175
 DB |||||
 QY 484 GACCTTCGCGGATCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 543
 DB |||||
 QY 1176 GATCTCTGCGGATCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1235
 DB |||||
 QY 544 GAGACGGTGCAGGACTGCAATGCTCAATCTATCTTCCCGGCGGACATGACGGTCACTG 603
 DB |||||
 QY 1236 GAGACGGTGCAGGACTGCAATGCTCAATCTATCTTCCCGGCGGACATGACGGTCACTG 1295
 DB |||||
 QY 604 GCTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 663
 DB |||||
 QY 1296 GCTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1355
 DB |||||
 QY 664 CGGATCCGACAGCTGCTGGGACATGCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 723
 DB |||||
 QY 1356 CGGATCCGACAGCTGCTGGGACATGCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1415
 DB |||||
 QY 724 CTCGCTTACTATTCATGCTGGGAACTGGGCTTAAGGTTTGTATGATGATGATGATGATG 783
 DB |||||
 QY 1416 CTTGCTTACTATTCATGCTGGGAACTGGGCTTAAGGTTTGTATGATGATGATGATGATG 1475
 DB |||||
 QY 784 GC 785
 DB ||
 QY 1476 GC 1477

RESULT 11
 ABK91424
 ID ABK91424 standard; DNA; 9605 BP.
 XX
 AC ABK91424;
 XX
 DT 15-NOV-2002 (first entry)
 XX
 DE Hepatitis C virus Con 1 isolate DNA mutant 1.
 XX
 KW HCV; ds; Con 1; adaptive mutation; liver failure; cirrhosis;
 KW hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;
 KW internal ribosome entry site; IRES; NS5A; HCV replication; mutant.
 XX
 OS Hepatitis C virus.
 OS Synthetic.
 XX
 XX Key Location/Qualifiers
 CDS 342..9374
 /tag= a
 /product= "HCV polyprotein"
 /note= "The polyprotein consists of the Core, E1, E2, P7,
 NS2, NS3, NS4A, NS4B, NS5A and NS5B proteins"
 FT mutation replace(3625,G)
 /tag= b
 XX
 PN WO200259321-A2.
 XX
 PD 01-AUG-2002.
 XX
 PP 16-JAN-2002; 2002WO-EP000526.
 XX
 PR 23-JAN-2001; 2001US-0263479P.
 XX
 PA (RICE-) IST RICERCH BIOL MOLECOLARE ANGELETTI.
 XX
 PI De Francesco R, Migliaccio G, Paonessa G;
 DR WPI; 2002-599793/64.
 XX
 PT New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV
 PT NS5 encoding region, or encephalomyocarditis virus (EMCV) internal
 PT ribosome entry site (IRES) region, useful in studying HCV replication and
 PT expression.
 XX
 PS Claim 9; Page; 69pp; English.
 XX
 CC The invention relates to nucleic acid molecules comprising altered HCV
 CC NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)
 CC internal ribosome entry site (IRES) region coding for one or more NS3,
 CC NS5A, or EMCV IRES mutations, respectively. The location of the mutations
 CC are detailed in the specification. Also included are (1) an expression
 CC vector comprising a nucleotide sequence coding for the altered nucleic
 CC acids, which is transcriptionally coupled to an exogenous promoter; (2) a
 CC recombinant cell human hepatoma cell comprising the altered nucleic acids
 CC ; (3) a recombinant cell produced by introducing into a human hepatoma
 CC cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)
 CC replicon enhanced cell or which containing a functional HCV replicon; (5)
 CC an HCV replicon enhanced cells made in the method; and (6) measuring the
 CC ability of a compound to affect HCV activity. The HCV replicons and HCV
 CC replicon enhanced cells are useful in studying HCV replication and
 CC expression, and HCV and host cell interactions, producing HCV RNA and
 CC proteins, and providing a system for measuring the ability of a compound
 CC to modulate one or more HCV activities e.g. to discover drugs which may
 CC treat HCV mediated diseases such as liver failure, cirrhosis and
 CC hepatocellular carcinoma. The present sequence is an HCV replicon Con 1
 CC mutant of the invention. Note: The present sequence is not shown in the
 CC specification but was created by the indexer using the HCV sequence
 CC appearing as ABK91411 and the information in Claim 9
 XX
 SQ Sequence 9605 BP; 1910 A; 2884 C; 2732 G; 2079 T; 0 U; 0 Other;
 Query Match 88.5%; Score 703.6; DB 6; Length 9605;

Db 816 GAGGACGGCGTGAACATATCAACAGGGAATCTGCCGGTGTCTCTTTCTATCTTCCTT 875
 QY 184 TTGGCTTTGTCTCTCTCTGACCGTTCCAGCTTCCGCTTATGAAGTGGCAAGTGTGCC 243
 Db 876 TTGGCTTTGTCTCTCTCTGACCGTTCCAGCTTCCGCTTATGAAGTGGCAAGTGTGCC 935
 QY 244 GGGATGATACATGTCACGAACGACTGCTTCCAACTCAAGCAATTTGTATGAGCGCAGCGAC 303
 Db 936 GGAAGTGTACCATGTCACGAACGACTGCTTCCAACTCAAGCAATTTGTATGAGCGCAGCGAC 995
 QY 304 ATGATCATGACACACCCCGGGTGGCTGCTTCCGCTTCCGAGGAGCAACTCTTCCGCTGC 363
 Db 996 ATGATCATGACATACCCCGGGTGGCTGCTTCCGCTTCCGAGGAGCAACTCTTCCGCTGC 1055
 QY 364 TGGGTAGCGCTCACCCCGGCTCGCAGCTAGGAGGAGCGGCTGCTTCCGCTTCCGAGGAGCAACTCTTCCGCTGC 423
 Db 1056 TGGGTAGCGCTCACCCCGGCTCGCAGCTAGGAGGAGCGGCTGCTTCCGCTTCCGAGGAGCAACTCTTCCGCTGC 1115
 QY 424 CGACGCCACATGCTGATTTGCTCTGTTGGGGCGGCTGCTTCTGTTCCGCTATGTAGTGGGG 483
 Db 1116 CGACGCCACATGCTGATTTGCTCTGTTGGGGCGGCTGCTTCTGTTCCGCTATGTAGTGGGA 1175
 QY 484 GACCTCTGCGAGCTGTCTCTGCTCTCCAGCTGTTCACCATCTCGCTTCCGCTTCCGAGGAT 543
 Db 1176 GATCTCTGCGAGCTGTCTCTGCTCTCCAGCTGTTCACCATCTCGCTTCCGCTTCCGAGGAT 1235
 QY 544 GAGACGCTGACGAGCTGCAATTTGCTCAATCTATCCCGGCGCACATAAGCGGTCAACGCTATG 603
 Db 1236 GAGACGCTGACGAGCTGCAATTTGCTCAATCTATCCCGGCGCACATAAGCGGTCAACGCTATG 1295
 QY 604 GCTTGGGATATGATGAACTGCTGCTTCAACGCGGCTTGTGTTGATGCGAGCTGTCTC 663
 Db 1296 GCTTGGGATATGATGAACTGCTGCTTCAACGCGGCTTGTGTTGATGCGAGTACTC 1355
 QY 664 CGATCCCAAGCTGTCTGACATGCTGCGGGGCGGCTTGTGTTGATGCGAGTACTC 723
 Db 1356 CGATCCCAAGCTGTCTGACATGCTGCGGGGCGGCTTGTGTTGATGCGAGTACTC 1415
 QY 724 CTCGCTTACTTATCCATGTTGGGAACTGCGGCTTGTGTTGATGCGAGTACTC 783
 Db 1416 CTCGCTTACTTATCCATGTTGGGAACTGCGGCTTGTGTTGATGCGAGTACTC 1475
 QY 784 GC 785
 Db 1476 GC 1477

RESULT 13
 ID ABK91432
 XX AC ABK91432;
 XX AC ABK91432;
 DT 15-NOV-2002 (first entry)
 XX Hepatitis C virus Con 1 isolate DNA mutant 9.
 XX Hepatitis C virus.
 XX Synthetic.
 XX Key
 XX CDS
 XX Location/Qualifiers
 342..9374
 /tag= a
 /product= "HCV polyprotein"
 /note= "The polyprotein consists of the Core, E1, E2, P7,
 NS2, NS3, NS4A, NS4B, NS5A and NS5B proteins"
 mutation
 replace(6936,6)
 /tag= b

PN WO200259321-A2.
 XX 01-AUG-2002.
 XX 16-JAN-2002; 2002MO-EP000526.
 XX 23-JAN-2001; 2001US-0263479P.
 XX (RICE-) IST RICERHE BIOL MOLECOLARE ANGELETTI.
 XX De Francesco R, Migliaccio G, Paonessa G;
 XX WPI; 2002-599793/64.
 XX New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV
 NS5 encoding region, or encephalomyocarditis virus (EMCV) internal
 ribosome entry site (IRES) region, useful in studying HCV replication and
 expression.
 Claim 9; Page; 69pp; English.
 CC The invention relates to nucleic acid molecules comprising altered HCV
 NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)
 internal ribosome entry site (IRES) region coding for one or more NS3,
 NS5A, or EMCV IRES mutations, respectively. The location of the mutations
 are detailed in the specification. Also included are (1) an expression
 vector comprising a nucleotide sequence coding for the altered nucleic
 acids, which is transcriptionally coupled to an exogenous promoter; (2) a
 recombinant cell human hepatoma cell comprising the altered nucleic acids
 ; (3) a recombinant cell produced by introducing into a human hepatoma
 cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)
 replicon enhanced cell or which containing a functional HCV replicon; (5)
 an HCV replicon enhanced cells made in the method; and (6) measuring the
 ability of a compound to affect HCV activity. The HCV replicons and HCV
 replicon enhanced cells are useful in studying HCV replication and
 expression, and HCV and host cell interactions, producing HCV RNA and
 proteins, and providing a system for measuring the ability of a compound
 to modulate one or more HCV activities e.g. to discover drugs which may
 treat HCV mediated diseases such as liver failure, cirrhosis and
 hepatocellular carcinoma. The present sequence is an HCV replicon Con 1
 mutant of the invention. Note: The present sequence is not shown in the
 specification but was created by the indexer using the HCV sequence
 appearing as ABK91411 and the information in Claim 9
 XX Sequence 9605 BP; 1911 A; 2883 C; 2732 G; 2079 T; 0 U; 0 Other;
 Query Match 88.5%; Score 703.6; DB 6; Length 9605;
 Best Local Similarity 93.7%; Pred. No. 4e-179;
 Matches 733; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
 QY 4 TTGGTAAAGCTCATCGATACCTTTACATGCGGCTTGGCGGCTGCGGCTTACATTCCG 63
 Db 696 TTGGTAAAGCTCATCGATACCTTTACATGCGGCTTGGCGGCTGCGGCTTACATTCCG 755
 QY 64 CTCGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 123
 Db 756 CTCGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 815
 QY 124 GAGACGCGGCTGAACTATGCAAGAGGAAATTTGCGGCTGCTCTTCTATCTTCCTC 183
 Db 816 GAGACGCGGCTGAACTATGCAAGAGGAAATTTGCGGCTGCTCTTCTATCTTCCTC 875
 QY 184 TTGGCTTTGTCTCTCTCTGACCGTTCCAGCTTCCGCTTATGAAGTGGCAAGTGTGCC 243
 Db 876 TTGGCTTTGTCTCTCTCTGACCGTTCCAGCTTCCGCTTATGAAGTGGCAAGTGTGCC 935
 QY 244 GGGATGATACATGTCACGAACGACTGCTTCCAACTCAAGCAATTTGTATGAGCGCAGCGAC 303
 Db 936 GGAAGTGTACCATGTCACGAACGACTGCTTCCAACTCAAGCAATTTGTATGAGCGCAGCGAC 995
 QY 304 ATGATCATGACACACCCCGGGTGGCTGCTTCCGCTTCCGAGGAGCAACTCTTCCGCTGC 363
 Db 996 ATGATCATGACATACCCCGGGTGGCTGCTTCCGCTTCCGAGGAGCAACTCTTCCGCTGC 1055

QY 364 TGGGTAGCGCTCACCCACGCTCGCAGCTAGGAGCGCAGCGTCCCGACACGACAATA 423
 Db 1056 TGGGTAGCGCTCACTCCACGCTCGCGGCCAGGAGCTAGCGTCCCGACTACGACGATA 1115
 QY 424 CGAGCCGACGCTGATTTGCTGTTGGGGGGGCTGCTTCTGTTCCGCTATGACGTGGGG 483
 Db 1116 CGAGCCGACGCTGATTTGCTGTTGGGGGGGCTGCTTCTGTTCCGCTATGACGTGGGA 1175
 QY 484 GACCTCTGGGATGCTGCTTCTGCTGCTCCAGCTGTTCCACCATCTCGCTCGCGGCAT 543
 Db 1176 GATCTCTGGGATGCTGCTTCTGCTGCTCCAGCTGTTCCACCATCTCGCTCGCGGCAC 1235
 QY 544 GAGACGCTGAGGACTGCAATGCTCAATCTATCCCGGCCACATACCGGCTCACCGTATG 603
 Db 1236 GAGACAGTACGAGCTGCAATGCTCAATATATCCCGGCCAGCTGACAGGTACCGTATG 1295
 QY 604 GCTTGGGATATGATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 663
 Db 1296 GCTTGGGATATGATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1355
 QY 664 CGGATCCACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 723
 Db 1356 CGGATCCACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1415
 QY 724 CTCGCTACTATTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 783
 Db 1416 CTGCTACTATTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1475
 QY 784 GC 785
 Db 1476 GC 1477

RESULT 14
 ID ABK91411
 AC ABK91411;
 XX
 DT 15-NOV-2002 (first entry)
 XX
 DE Hepatitis C virus Con 1 isolate DNA.
 XX
 KW HCV; ds; Con 1; adaptive mutation; liver failure; cirrhosis;
 KW hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;
 KW internal ribosome entry site; IRES; NS5A; HCV replication.
 XX
 OS Hepatitis C virus.
 XX
 FH Location/Qualifiers
 FT CDS 342..9374
 FT /*tag= a
 FT /product= "HCV polyprotein"
 FT /note= "The polyprotein consists of the Core, E1, E2, E7,
 NS2, NS3, NS4A, NS4B, NS5A and NS5B proteins"
 XX
 EN WO200259321-A2.
 XX
 PD 01-AUG-2002.
 XX
 PF 16-JAN-2002; 2002WO-EP000526.
 XX
 PR 23-JAN-2001; 2001US-0263479P.
 XX
 XX (RICE-) IST RICERHE BIOL MOLECOLARE ANGELETTI.
 PA De Francesco R, Migliaccio G, Paonessa G;
 XX
 XX WPI; 2002-599793/64.
 DR P-PSDB; ABG32451.
 XX
 PT New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV

PT NS5 encoding region, or encephalomyocarditis virus (EMCV) internal
 PT ribosome entry site (IRES) region, useful in studying HCV replication and
 PT expression.
 XX Claim 9; Page 36-39; 69pp; English.
 PS
 CC The invention relates to nucleic acid molecules comprising altered HCV
 CC NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)
 CC internal ribosome entry site (IRES) region coding for one or more NS3,
 CC NS5A, or EMCV IRES mutations, respectively. The location of the mutations
 CC are detailed in the specification. Also included are (1) an expression
 CC vector comprising a nucleotide sequence coding for the altered nucleic
 CC acids, which is transcriptionally coupled to an exogenous promoter; (2) a
 CC recombinant cell human hepatoma cell comprising the altered nucleic acids
 CC ; (3) a recombinant cell produced by introducing into a human hepatoma
 CC cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)
 CC replicon enhanced cell or which containing a functional HCV replicon; (5)
 CC an HCV replicon enhanced cells made in the method; and (6) measuring the
 CC ability of a compound to affect HCV activity. The HCV replicons and HCV
 CC replicon enhanced cells are useful in studying HCV replication and
 CC expression, and HCV and host cell interactions, producing HCV RNA and
 CC proteins, and providing a system for measuring the ability of a compound
 CC to modulate one or more HCV activities e.g. to discover drugs which may
 CC treat HCV mediated diseases such as liver failure, cirrhosis and
 CC hepatocellular carcinoma. The present sequence is the HCV replicon Con 1,
 CC used as a basis for the adaptive mutations of the invention
 XX
 SQ Sequence 9605 BP; 1910 A; 2883 C; 2733 G; 2079 T; 0 U; 0 Other;

Query Match 88.5%; Score 703.6; DB 6; Length 9605;
 Best Local Similarity 93.7%; Pred. No. 4e-179;
 Matches 733; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
 QY 4 TTGGGTAAAGTCAATGATACCTTACATGGGCTTCGCGGACCTCTCGGGTACATTCOG 63
 Db 696 TTGGGTAAAGTCAATGATACCTTACATGGGCTTCGCGGATCTCATAGGGTACATTCOG 755
 QY 64 CTCGTGCGCGCCCCCTAGCGGGCGCTGCGAGGGCCCTGCGCGATGCGGCTTCG 123
 Db 756 CTCGTGCGCGCCCCCTAGCGGGCGCTGCGAGGGCCCTGCGCGATGCGGCTTCG 815
 QY 124 GAGGACGGGTGAATATATGCAACAGGGAATTTGCCGGTTCCTCTTCTATCTTCCTC 183
 Db 816 GAGGACGGGTGAATATATGCAACAGGGAATCTGCCGGTTCCTCTTCTATCTTCCTC 875
 QY 184 TTGGCTTTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 243
 Db 876 TTGGCTTTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 935
 QY 244 GGGATGTACCATGTCAAGAACGACCTGCTCCAACTCAAGCATTTGTATGAGCGCGGAC 303
 Db 936 GGGATGTACCATGTCAAGAACGACCTGCTCCAACTCAAGCATTTGTATGAGCGCGGAC 995
 QY 304 ATGATCATGACACACCCCGGGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 363
 Db 996 ATGATCATGACACACCCCGGGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1055
 QY 364 TGGGTAGCGCTCACCCCAAGCTGCGAGTACGAGACCGCAGCGTTCCTCCACGACATA 423
 Db 1056 TGGGTAGCGCTCACCTCCCAAGCTGCGAGTACGAGACCGTTCCTCCACGACATA 1115
 QY 424 CGAGCCGACGCTGATTTGCTGTTGGGGGGGCTGCTTCTGTTCCGCTATGACGTGGGG 483
 Db 1116 CGAGCCGACGCTGATTTGCTGTTGGGGGGGCTGCTTCTGTTCCGCTATGACGTGGGA 1175
 QY 484 GACCTCTGGGATGCTGCTTCTGCTGCTCCAGCTGTTCCACCATCTCGCTCGCGGCAT 543
 Db 1176 GATCTCTGGGATGCTGCTTCTGCTGCTCCAGCTGTTCCACCATCTCGCTCGCGGCAC 1235
 QY 544 GAGACGCTGAGGACTGCAATGCTCAATCTATCCCGGCCACATACCGGCTCACCGTATG 603
 Db 1236 GAGACAGTACGAGCTGCAATGCTCAATATATCCCGGCCAGCTGACAGGTACCGTATG 1295

Db ||
 1476 GC 1477

Search completed: June 16, 2004, 09:36:53
Job time : 301.119 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2004, 09:12:53 ; Search time 2137.22 Seconds
(without alignments)
11108.065 Million cell updates/sec

Title: US-09-899-303A-5

Perfect score: 795

Sequence: 1 ATGTTGGTAAAGTCATCGA.....TACTCTTGTCTCCCTAATAG 795

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estnu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_estl:*

9: gb_est1:*

10: gb_est2:*

11: gb_est3:*

12: gb_est4:*

13: gb_est5:*

14: gb_est6:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_phg:*

27: em_gss_vrl:*

28: gb_gss1:*

29: gb_gss2:*

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	115.6	14.5	488	9	AV755731
c 2	91.6	11.5	492	9	AV758366
c 3	47.2	5.9	1201	13	BX356664
4	43.4	5.5	925	29	CNS0091P

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

C	5	43	5.4	886	14	CK209050
	6	42.8	5.4	423	9	AA459034
C	7	41.6	5.2	1201	9	AL513886
	8	41.2	5.2	1195	28	BL10902
	9	40.6	5.1	402	9	AV392783
	10	40.6	5.1	551	9	AV392165
	11	40.6	5.1	552	12	BI996341
	12	40.6	5.1	584	12	BI727879
	13	40	5.0	525	10	BE337089
	14	40	5.0	534	14	CF846043
	15	40	5.0	671	12	BI723733
C	16	39.8	5.0	608	14	CB640103
C	17	39.8	5.0	738	14	CB668031
	18	39.8	5.0	969	12	BM017656
C	19	39.6	5.0	675	12	BQ045314
	20	39.2	4.9	571	12	BM692316
	21	39.2	4.9	645	29	CNS01213
C	22	39.2	4.9	1102	13	BQ069967
	23	39	4.9	359	12	BQ252669
	24	39	4.9	375	12	BQ246716
	25	39	4.9	621	14	CA816001
	26	39	4.9	624	12	BI723734
C	27	39	4.9	739	29	CG239174
C	28	39	4.9	744	29	CC605518
C	29	39	4.9	828	29	CG381985
	30	39	4.9	840	28	CC335916
	31	39	4.9	841	29	CC605508
	32	39	4.9	873	14	CD446071
C	33	39	4.9	1029	29	CC625582
	34	38.8	4.9	590	14	CA659369
	35	38.6	4.9	693	13	CA160850
	36	38.4	4.8	702	14	CD432549
C	37	38.4	4.8	740	12	BQ536071
	38	38.4	4.8	970	29	CNS010C9
C	39	38.4	4.8	987	29	CNS015VX
C	40	38.2	4.8	1148	14	CK208676
C	41	38.2	4.8	1270	12	BQ968359
	42	38	4.8	354	14	CB966525
C	43	38	4.8	742	13	BQ752673
	44	37.8	4.8	435	13	C72860
C	45	37.8	4.8	533	28	CC010084

ALIGNMENTS

RESULT 1	AV755731/c	488 bp	mRNA	linear	EST 19-OCT-2000
LOCUS	AV755731	BM	Homo sapiens	cDNA clone BMFARB03 5', mRNA sequence.	
DEFINITION	AV755731				
ACCESSION	AV755731				
VERSION	AV755731.1	GI:10913579			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	Gu, J., Zhao, M., Huang, Q., Xu, X., Li, Y., Peng, Y., Song, H., Xiao, H., Gu, Y., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z., Zeng, L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G., Yang, Y., Gao, G., Wang, Z., Zhang, Q., Chen, S., Han, Z. and Chen, Z.				
TITLE	Homo sapiens cDNA BM clones				
JOURNAL	Unpublished (2000)				
COMMENT	Contact: Zeguang Han Chinese National Human Genome Center at Shanghai 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China Tel: 86-21-50801919(ex.45) Fax: 86-21-50801922 Email: hanzg@chgc.sh.cn This clone is available at CHGC in Shanghai.				

sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN		FEATURES		source	
Query Match		5.9%; Score 47.2; DB 13; Length 1201;		1..925	
Best Local Similarity		10.7%; Pred. No. 6.1;		/organism="Drosophila melanogaster"	
Matches	57; Conservative 250; Mismatches 227; Indels 0; Gaps 0;				/mol_type="genomic DNA"
					/db_xref="taxon:7227"
					/clone="BACR19D16"
					/clone_lib="RPCI-98"
					/note="end : TET3"
ORIGIN		Query Match		5.5%; Score 43.4; DB 29; Length 925;	
		Best Local Similarity	16.1%; Pred. No. 39;		
		Matches	70; Conservative 166; Mismatches 196; Indels 4; Gaps 1;		
Qy	23	CCCTTACATCGGCTTCGCCACCTCGTGGGTACATTCGCTCGTGGGCCCCCTAG	82		
Db	665	SSBTBSTSSSSSSSBTTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSS	724		
Qy	83	GGGCGCTGCGCAGGCGCTCGCGATCGCGTCCGGTTCGAGAGCGCGCTGAATG	142		
Db	725	SSSTBSSTTCTTTKBSSTTBSSTTBSSTTBSSTTBSSTTBSSTTBSSTTBSST	784		
Qy	143	CAACAGGAATTCGCCGGTGTCTTTCTCTATCTCTCTCTCTCTCTCTCTCTCTCT	202		
Db	785	SCSCTSSSTTSSBMTSSSBVTCTSSSSSSSBTTSSSTTSSSTTSSSTTSSST	844		
Qy	203	TGACCGTTCAGCTTCGCTTATGAGTGCACAGTGCAGGTGCGGATGTACATCTCA	262		
Db	845	SSTWTBTCTSTTNTSYSSSTTBSSTTBSSTTBSSTTBSSTTBSSTTBSSTTBS	904		
Qy	263	ACGACTCTCAACTCAAGCATCTGTATGAGGCGGACATCATCATGCACACCCCG	322		
Db	905	TTTBSSSSSSSSTTBSSTTBSSTTBSSTTBSSTTBSSTTBSSTTBSSTTBSST	964		
Qy	323	GGTGGCGGCTCGTTCGGAGAACACTTCTCCGCTGCTGCGTACGCGCTACCCCA	382		
Db	965	SBSSSSSSSTTBSSTTBSSTTBSSTTBSSTTBSSTTBSSTTBSSTTBSSTTBS	1024		
Qy	383	CGTCCGAGCTAGAACCGCAGCTCCACCACGACATACAGACGCGCATCTCATTC	442		
Db	1025	SSSTTBSSTSSNNKSSSSSSSSSSSSSSSSSSSSSSSTTBSSTTBSSTTBSST	1084		
Qy	443	TGCTTGGGCGGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	502		
Db	1085	TSSBSSSSSTTBSSTTBSSTTBSSTTBSSTTBSSTTBSSTTBSSTTBSSTTBS	1144		
Qy	503	TCTGCTCCAGCTGTTCCACATCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCG	556		
Db	1145	SSTTSTTTTBTATSTTTTBSBSSSSSSSSSSSTTBSSTTBSSTTBSSTTBSST	1198		
RESULT 4		CNS0091P		925 bp DNA linear GSS 03-JUN-1999	
LOCUS	CNS0091P				
DEFINITION	Drosophila melanogaster genome survey sequence TET3 end of BAC #				
	BACR19D16 of RPCI-98 library from Drosophila melanogaster (fruit				
	fly), genomic survey sequence.				
ACCESSION	AL053013				
VERSION	AL053013.1				
KEYWORDS	GSS.				
SOURCE	Drosophila melanogaster (fruit fly)				
	Drosophila melanogaster				
	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;				
	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;				
	Ephydroidea; Drosophilidae; Drosophila.				
	1 (bases 1 to 925)				
REFERENCE	Genoscope.				
AUTHORS	Direct Submission				
TITLE	Submitted (02-JUN-1999) Genoscope - Centre National de Sequençage :				
JOURNAL	BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr				
	- Web : www.genoscope.cns.fr)				
	Determination of this BAC-end sequence was carried out as part of a				
	collaboration with the Berkeley Drosophila Genome Project (BDGP).				
	The BDGP is constructing a physical map of the Drosophila				
	melanogaster genome using these BACs. For further information				
	please see http://www.fruitfly.org The BDGP Drosophila				
	melanogaster BAC library was prepared by Kazutoyo Osogawa and				
	Aaron Mamoser in Pieter de Jong's laboratory in the Department of				
	Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,				
	NY. The library is named RPCI-98 and was constructed by partial				
	unpublished (2003)				
RESULT 5		CK209050/c		886 bp mRNA linear EST 08-DEC-2003	
LOCUS	CK209050				
DEFINITION	FGAS020786 Triticum aestivum FGAS: Library 5 GATE 7 Triticum				
	aestivum cDNA, mRNA sequence.				
ACCESSION	CK209050				
VERSION	CK209050.1				
KEYWORDS	EST.				
SOURCE	Triticum aestivum (bread wheat)				
	Triticum aestivum				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;				
	Pooideae; Triticeae; Triticum.				
	1 (bases 1 to 886)				
REFERENCE	Allard, F., Crosby, W.L., Danyluk, J., Eudes, F., Frick, M., Gaudet, D.,				
AUTHORS	Genswein, B., Graf, R., Gulick, P., Hrycan, L.D., Laroche, A.,				
	Links, M.G., McCarthy, E.L., Monroy, A., Muzak, I., Nilsson, D.,				
	Penniket, C., Roach, J.L. and Sarhan, F.				
	Functional Genomics of Abiotic Stress in Wheat and Canola Crops				
	Unpublished (2003)				

EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's p1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

COMMENT

Contact: Wm L Crosby
Bioinformatics
University of Saskatchewan, Department of Computer Science
1C101 Engineering Building, 57 Campus Drive, Saskatoon,
Saskatchewan, S7N 5A9, Canada
Tel: 306 966 1769
Fax: 306 966 2033
Email: fgas_est@cs.usask.ca
This sequence is the direct result of the Base calling software
phred (default parameters). It is the raw base calls. To aid in the
identification of the high quality insert the software Lucy
(default parameters) has been run on this sequence. Lucy identified
the region [65,694].

FEATURES

source

Location/Qualifiers

1. 886

/organism="Triticum aestivum"

/mol_type="mRNA"

/db_xref="taxon:4565"

/clone_lib="Triticum aestivum FGAS: Library 5 GATE 7"

/note="Vector: pCMV.SPORT6; Crown and developmental stages

of spike formation in wheat cultivar Norstar. 4 mRNA

populations were combined before constructing the library.

The first mRNA population is from lcm crown sections after

30 days of cold acclimation. The second is from lcm crown

sections after 11 days of deacclimation (before

deacclimation plants were fully vernalized for 49 days).

The third is from different developmental stages of spike

formation (5 to 50mm) that still have not emerged from the

leaf (dissection required). The last is from different

developmental stages of spike and seed formation after

having emerged from the leaf (visible). First strand

synthesis in this library was done in the presence of

methylated dCTP thereby protecting from internal cleavage

with NotI."

ORIGIN

Query Match 5.4%; Score 43; DB 14; Length 886;
Best Local Similarity 54.8%; Pred. No. 48;
Matches 85; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

Qy 11 AGGTGATGATACCTTATGATGGCTTGGCGGACCTGCGGGTACATTCGCTGCTG 70
Db 669 AGGCTTCGACGACACAGCGGGGCGCGGAGACATCGCGGCGACGCGCTCCGCGACA 610
Qy 71 GCGGCCCCCTAGGGGGCGCTGCGAGGCGCTGGCGGCGAGCTGCGGGTTCGAGGAGCG 130
Db 609 CGGCGGCGATCGCTGCGGCGCGCGCGGAGCTGTACGGCTCCAGGCTCTGCGGCGACG 550
Qy 131 CGGTGACATGATGACACAGCGGAATTTGCCCGGTTGC 165
Db 549 CGGTGACAGGACGACCGCGGAGACGTGACCGGTTTC 515

RESULT 6

AA459034

LOCUS

DEFINITION

AA459034

VERSION

AA459034.1

KEYWORDS

EST.

SOURCE

Homo sapiens

ORGANISM

Homo sapiens

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AA459034 423 bp mRNA linear EST 13-AUG-1997
a26b11.r1 NCI CGAP_GCB1 Homo sapiens cDNA clone IMAGE:814365 5',
mRNA sequence.
AA459034 1 GI:2183941
AA459034.1
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 423)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov

This clone is available royalty-free through LML; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 382.

FEATURES

source

Location/Qualifiers

1. 423

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="GDB:6032522"

/db_xref="taxon:9606"

/clone="IMAGE:814365"

/tissue_type="germinal center B cell"

/lab_host="DH10B"

/clone_lib="NCI CGAP_GCB1"

/note="Vector: pT7T3D-Pac (Pharmacia) with a modified

polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA

was prepared from human tonsillar cells enriched for

germinal center B cells by flow sorting (CD20+, IgD-),

provided by Dr. Louis M. Staudt (NCI), Dr. David Allman

(NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was

primed with a Not I - oligo(dT) primer

[5'-TGTTACCAATCTGAAGTGGGAGCGCGCTCATTTTCTTTT-3'

]. Double-stranded cDNA was ligated to Eco RI adaptors

(Pharmacia), digested with Not I and cloned into the Not I

and Eco RI sites of the modified pT7T3 vector. Library

went through one round of normalization, and was

constructed by Bento Soares and M. Fatima Bonaldo."

Durham, NC 27708-1000
Tel: 919 613 8159
Fax: 919 613 8177
Email: chauser@duke.edu

FEATURES Location/Qualifiers

/organism="Chlamydomonas reinhardtii"
 /mol_type="mRNA"
 /strain="CC-1690 wild type mt+ 21gr"
 /db_xref="taxon:3055"
 /clone.lib="C. reinhardtii CC-1690, Stress II
 (normalized), Lambda Zap II"
 /note="vector: pBluescript II SK-; Site 1: EcoRI;
 XhoI; Stress condition II library, constructed
 by Davies and Jeffrey McDermott, combines cDNAs
 from cells grown to mid-log phase in TAP (NH₄⁺ -
 depleted) and shifted to TAP - NO₃⁻ (24hrs); H2
 production conditions (0, 12hr, 24hr) see Melis et al.,
 Phys. 122: 127-135; TAP + H2O₂ (1, 12, 24 hr)
 sorbitol (1, 2, 6, 24 hr); TAP + Cd (1, 2, 6,
 24 hr). PolyA mRNA was purified from each
 sample, pooled and directionally cloned and
 synthesized. The cDNA was directionally cloned
 into Zip II (Stratagene) in the EcoRI (5') and
 XhoI (3') sites. pBluescript II SK- plasmids were
 excised from lambda Zap clones by superinfection
 with ExAmp (Stratagene) phage. The library was
 normalized using the method described in
 Bionaldo et al., (1996) C. R. Acad. Sci. Paris
 Research 5: 791-806."

ORIGIN

Query Match	5.0%;	Score 40;	DB 12;	Length 671;
Best Local Similarity	47.0%;	Pred. No. 2e+02;		
Matches 124;	Conservative 0;	Mismatches 140;	Indels 0;	Gaps 0;
Qy	261	GAACGACTGCTCAACTCAAGCATTTGTTATGAGCGAGCGGACATGATCATGCACACCCC	320	
Db	353	GCACCGCTTCAACAACACCCACCGTCTGTAATTCGCCGCCCTACAAGTACCTGCTGTCAT	412	
Qy	321	CGGTCGTCGCTCGCTTCGGGAGAACAACTCTTCCCGCTGCTGGGTACGGCTGCACCCC	380	
Db	413	GGAGGCAAGACCTTCCCACTCGGAGGTCTACGAGTACTGCGACTGGGGCAACCTGTCCAA	472	
Qy	381	CACGCTCGCAGCTAGGAACGCCACGCTGCCCAACACAGCAATACAGCCACGTCGATTT	440	
Db	473	CGCGCTCAAAAACAATCTTATGATCCCCAACCCCGTCATCGCGCGCGCGGGCGGC	532	
Qy	441	GCTCGTTGGCGCGCTGCTTTTCTGTTCCGCTATGTACGTGGGGACCTTCGCGGATCTGT	500	
Db	533	GGCGCAGCGCGCGCGCGCAGAGCTAGCGGAGCGCGCGCAGCAGAGCCATGAGGT	592	
Qy	501	CTTCTCTGCTCCGAGCTGTTTAC	524	
Db	593	CAACATGCGCACCTGCTGCTCAC	616	

Search completed: June 16, 2004, 13:27:11
Job time : 2140.22 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2004, 09:14:33 ; Search time 57.3062 Seconds
(without alignments)
7698.741 Million cell updates/sec

Title: US-09-899-303A-5
Perfect score: 795
Sequence: 1 ATCTGGTGAAGTCATCGA.....TACTCTTTGCTCCTCAATAG 795

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/2/ina/5A COMB.seq: *
2: /cgn2_6/ptodata/2/ina/5B COMB.seq: *
3: /cgn2_6/ptodata/2/ina/6A COMB.seq: *
4: /cgn2_6/ptodata/2/ina/6B COMB.seq: *
5: /cgn2_6/ptodata/2/ina/PCTUS COMB.seq: *
6: /cgn2_6/ptodata/2/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	795	100.0	795	3	US-08-612-973-5
2	795	100.0	795	3	US-08-927-597-5
3	778.8	98.0	2082	3	US-08-612-973-47
4	778.8	98.0	2082	3	US-08-927-597-47
5	778.8	98.0	2433	3	US-08-612-973-49
6	778.8	98.0	2433	3	US-08-927-597-49
7	703.6	88.5	11076	4	US-09-539-601-1
8	703.6	88.5	11076	4	US-09-539-601-19
9	703.6	88.5	11076	4	US-09-539-601-25
10	703.6	88.5	11076	4	US-09-539-601-31
11	698.8	87.9	1539	2	US-08-470-426B-17
12	698.8	87.9	1863	2	US-08-470-426B-14
13	697.2	87.7	2116	3	US-08-191-160-21
14	695.6	87.5	932	1	US-08-081-072-15
15	695.6	87.5	932	1	US-08-449-093A-15
16	694	87.3	9595	3	US-09-014-416-4
17	694	87.3	9599	3	US-09-014-416-6
18	685.4	86.2	1167	1	US-08-324-977-9
19	685.4	86.2	1167	2	US-08-384-616-9
20	685.4	86.2	1167	2	US-08-904-686A-9
21	685.4	86.2	1167	3	US-09-315-850-9
22	685.4	86.2	1499	1	US-08-324-977-3
23	685.4	86.2	1499	2	US-08-384-616-3
24	685.4	86.2	1499	2	US-08-904-686A-3
25	685.4	86.2	1499	3	US-09-315-850-3
26	685.4	86.2	6039	1	US-08-324-977-11
27	685.4	86.2	6039	2	US-08-384-616-11

28	685.4	86.2	6039	2	US-08-904-686A-11	Sequence 11, Appl
29	685.4	86.2	6039	3	US-09-315-850-11	Sequence 11, Appl
30	685.4	86.2	9030	1	US-08-324-977-13	Sequence 13, Appl
31	685.4	86.2	9030	2	US-08-384-616-13	Sequence 13, Appl
32	685.4	86.2	9030	2	US-08-904-686A-13	Sequence 13, Appl
33	685.4	86.2	9030	3	US-09-315-850-13	Sequence 13, Appl
34	685.4	86.2	9416	1	US-08-324-977-1	Sequence 1, Appl
35	685.4	86.2	9416	2	US-08-384-616-1	Sequence 1, Appl
36	685.4	86.2	9416	2	US-08-904-686A-1	Sequence 1, Appl
37	685.4	86.2	9416	3	US-09-315-850-1	Sequence 1, Appl
38	685.4	86.2	9416	4	US-08-823-895A-27	Sequence 27, Appl
39	684.4	86.1	9472	4	US-08-150-204E-96	Sequence 96, Appl
40	641	80.6	723	3	US-08-612-973-21	Sequence 21, Appl
41	641	80.6	723	3	US-08-927-597-21	Sequence 21, Appl
42	640.8	80.6	742	1	US-08-081-072-18	Sequence 18, Appl
43	640.8	80.6	742	1	US-08-449-093A-18	Sequence 18, Appl
44	628.2	79.0	642	3	US-08-612-973-3	Sequence 3, Appl
45	628.2	79.0	642	3	US-08-927-597-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-08-612-973-5
; Sequence 5, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHUYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,973
; FILING DATE: 11-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 795 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..792
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..789

US-08-612-973-5

Query Match 100.0%; Score 795; DB 3; Length 795;
Best Local Similarity 100.0%; Pred. No. 9.8e-201;
Matches 795; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 1 ATGTTGGTAAAGTCAATGATACCTTACATGCGGCTTCGCGGACCTCGTGGGGTACATT 60
Db 1 ATGTTGGTAAAGTCAATGATACCTTACATGCGGCTTCGCGGACCTCGTGGGGTACATT 60

Qy 61 CCGCTCGTGGCGGCGCCCTTAGAGGGCGCTGCAGGGCCCTGCGGCATGCGGTCCGGGT 120
Db 61 CCGCTCGTGGCGGCGCCCTTAGAGGGCGCTGCAGGGCCCTGCGGCATGCGGTCCGGGT 120

Qy 121 CTGAGGACGGGTGAATCATGCAACAGGGAATTTGCGCGGTGCTCTTCTCTATCTTC 180
Db 121 CTGAGGACGGGTGAATCATGCAACAGGGAATTTGCGCGGTGCTCTTCTCTATCTTC 180

Qy 181 CTCTTGGCTTTGCTGCTGCTGCTGACCGTTCAGCTTCGGCTTATGAAGTGGCAACGTG 240
Db 181 CTCTTGGCTTTGCTGCTGCTGCTGACCGTTCAGCTTCGGCTTATGAAGTGGCAACGTG 240

Qy 241 TCCGGGATGATACCATGTCAGAACGCTGCTCCAACTCAAGCATTTGTATGAGGACGG 300
Db 241 TCCGGGATGATACCATGTCAGAACGCTGCTCCAACTCAAGCATTTGTATGAGGACGG 300

Qy 301 GACATGATCATGCACACCCCGGGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
Db 301 GACATGATCATGCACACCCCGGGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360

Qy 361 TGCTGGGTAGCGTCAACCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
Db 361 TGCTGGGTAGCGTCAACCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420

Qy 421 ATAGACCCACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
Db 421 ATAGACCCACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480

Qy 481 GGGGACCTCTGGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
Db 481 GGGGACCTCTGGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540

Qy 541 CATGAGCGGTGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
Db 541 CATGAGCGGTGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600

Qy 601 ATGGCTGGGATATCATGATGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Db 601 ATGGCTGGGATATCATGATGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660

Qy 661 CTCGGATCCCAACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
Db 661 CTCGGATCCCAACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720

Qy 721 GGTCTGGCTACTATTCCATGCTGGGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db 721 GGTCTGGCTACTATTCCATGCTGGGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780

Qy 781 TTTGCTCCCTAATAG 795
Db 781 TTTGCTCCCTAATAG 795
```

RESULT 2

US-08-927-597-5
; Sequence 5, Application US/08927597
; Patent No. 624503
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE

TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
NUMBER OF SEQUENCES: 111

CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/927,597
FILING DATE:
CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/612,973
FILING DATE: 11-MAR-1996

ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1487-10

TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100

INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 795 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:

NAME/KEY: CDS
LOCATION: 1..792
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 1..789

US-08-927-597-5

Query Match 100.0%; Score 795; DB 3; Length 795;
Best Local Similarity 100.0%; Pred. No. 9.8e-201;
Matches 795; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 1 ATGTTGGTAAAGTCAATGATACCTTACATGCGGCTTCGCGGACCTCGTGGGGTACATT 60
Db 1 ATGTTGGTAAAGTCAATGATACCTTACATGCGGCTTCGCGGACCTCGTGGGGTACATT 60

Qy 61 CCGCTCGTGGCGGCGCCCTTAGAGGGCGCTGCAGGGCCCTGCGGCATGCGGTCCGGGT 120
Db 61 CCGCTCGTGGCGGCGCCCTTAGAGGGCGCTGCAGGGCCCTGCGGCATGCGGTCCGGGT 120

Qy 121 CTGAGGACGGGTGAATCATGCAACAGGGAATTTGCGCGGTGCTCTTCTCTATCTTC 180
Db 121 CTGAGGACGGGTGAATCATGCAACAGGGAATTTGCGCGGTGCTCTTCTCTATCTTC 180

Qy 181 CTCTTGGCTTTGCTGCTGCTGCTGACCGTTCAGCTTCGGCTTATGAAGTGGCAACGTG 240
Db 181 CTCTTGGCTTTGCTGCTGCTGCTGACCGTTCAGCTTCGGCTTATGAAGTGGCAACGTG 240

Qy 241 TCCGGGATGATACCATGTCAGAACGCTGCTCCAACTCAAGCATTTGTATGAGGACGG 300
Db 241 TCCGGGATGATACCATGTCAGAACGCTGCTCCAACTCAAGCATTTGTATGAGGACGG 300

Qy 301 GACATGATCATGCACACCCCGGGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
Db 301 GACATGATCATGCACACCCCGGGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
```


RESULT 4

US-08-927-597-47
; Sequence 47, Application US/08927597
; Patent No. 6245503
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/927,597
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/612,973
; FILING DATE: 11-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2082 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2079
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..2076
; US-08-927-597-47
Query Match 98.0%; Score 778.8; DB 3; Length 2082;
Best Local Similarity 99.7%; Pred. No. 2.5e-196;
Matches 780; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 4 TTGGGTAAGGTGATCGATACCTTACATCGGGCTTCGCCGACTCGTGGGGTACATTCGG 63
DB 4 TTGGGTAAGGTGATCGATACCTTACATCGGGCTTCGCCGACTCGTGGGGTACATTCGG 63
QY 64 CTCGTGGGGCCCCCTAGGGGGCGCTGCAGGGCCCTGGCGCATGGCTCGGGTTCG 123
DB 64 CTCGTGGGGCCCCCTAGGGGGCGCTGCAGGGCCCTGGCGCATGGCTCGGGTTCG 123
QY 124 GAGGACGGGTGAACATATGCAACAGGGAAATTTGCCCGTTGCTTTCTCTATCTTCCTC 183
DB 124 GAGGACGGGTGAACATATGCAACAGGGAAATTTGCCCGTTGCTTTCTCTATCTTCCTC 183
QY 184 TTGGGCTTGTCTGCTGCTGACCGGTTCCAGCTTCGCGCTTATGAAGTGGCAACGTGTC 243

DB 184 TTGGCTTTGCTGTCTGCTGCTGACCGGTTCCAGCTTCGCGCTTATGAAGTGGCAACGTGTC 243
QY 244 GGGATGTACCATGTCAAGAAAGACTGCTCCAACTCAAGCATTTGTGTATGAGGACGCGGAC 303
DB 244 GGGATGTACCATGTCAAGAAAGACTGCTCCAACTCAAGCATTTGTGTATGAGGACGCGGAC 303
QY 304 ATGATATGCAACACCCCGGGTGGTGGCTGCGGTTCGGGAGAAACAATCTTTCCCGCTGC 363
DB 304 ATGATATGCAACACCCCGGGTGGTGGCTGCGGTTCGGGAGAAACAATCTTTCCCGCTGC 363
QY 364 TGGGTAGCGCTCAACCCCAAGCTGCGAGTAGGAAGCGCAGCGTCCCAACACGACAAATA 423
DB 364 TGGGTAGCGCTCAACCCCAAGCTGCGAGTAGGAAGCGCAGCGTCCCAACACGACAAATA 423
QY 424 CGACGCCACGCTCGATTGCTGCTTGGGGGGGCTGCTTTCTGTTCCGCTATGTACGTGGGG 483
DB 424 CGACGCCACGCTCGATTGCTGCTTGGGGGGGCTGCTTTCTGTTCCGCTATGTACGTGGGG 483
QY 484 GACCTCTGGGATCTGTCTTCTGCTGCTCCAGCTGTTTCAACATCTCGCTCGCGGCTAT 543
DB 484 GACCTCTGGGATCTGTCTTCTGCTGCTCCAGCTGTTTCAACATCTCGCTCGCGGCTAT 543
QY 544 GAGCGGTGCGAGCTGCAATTTGCTCAATCTATCCCGGCCACATACCGGTTCACCGTATG 603
DB 544 GAGCGGTGCGAGCTGCAATTTGCTCAATCTATCCCGGCCACATACCGGTTCACCGTATG 603
QY 604 GCTTGGGATATGATGATGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 663
DB 604 GCTTGGGATATGATGATGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 663
QY 664 CGGATCCCAAGCTGCTGCGACATGCTGGCGGGGGCCCATTTGGGAGTCTTGGCGGGT 723
DB 664 CGGATCCCAAGCTGCTGCGACATGCTGGCGGGGGCCCATTTGGGAGTCTTGGCGGGT 723
QY 724 CTCGCTTACTATTTCCATGCTGGGGAACTGCGGCTTAAGTTTTCATTTGTGATGCTACTCTT 783
DB 724 CTCGCTTACTATTTCCATGCTGGGGAACTGCGGCTTAAGTTTTCATTTGTGATGCTACTCTT 783
QY 784 GC 785
DB 784 GC 785
RESULT 5
US-08-612-973-49
; Sequence 49, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,973
; FILING DATE: 11-MAR-1996
; CLASSIFICATION: 435
; NAME: BYRNE, THOMAS E.

```

/ REGISTRATION NUMBER: 32,205
/ REFERENCE/DOCKET NUMBER: 1487-10
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (703) 816-4000
/ TELEFAX: (703) 816-4100
/ INFORMATION FOR SEQ ID NO: 49:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 2433 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: cDNA
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 1..2430
/ FEATURE:
/ NAME/KEY: mat_peptide
/ LOCATION: 1..2427
/
US-08-612-973-49
Query Match 98.0%; Score 778.8; DB 3; Length 2433;
Best Local Similarity 99.7%; Pred. No. 2.7e-196;
Matches 780; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 TTGGGTAAGGTCAATGATACCCCTTACATGCGGCTTCGCCGACCTCGTGGGGTACATTCGG 63
DB |||||
DB 355 TTGGGTAAGGTCAATGATACCCCTTACATGCGGCTTCGCCGACCTCGTGGGGTACATTCGG 414
QY 64 CTGCTCGGCGCCCCCTAGGGGCGCTCCAGGCGCCCTGCGCATGCGCGTCCGGGTTCTG 123
DB |||||
DB 415 CTGCTCGGCGCCCCCTAGGGGCGCTCCAGGCGCCCTGCGCATGCGCGTCCGGGTTCTG 474
QY 124 GAGGACGGGTGAATATGATCAACAGGGAATTCGCCGCTGCTCTTCTATCTTCCTC 183
DB |||||
DB 475 GAGGACGGGTGAATATGATCAACAGGGAATTCGCCGCTGCTCTTCTATCTTCCTC 534
QY 184 TTGGCTTTGCTGCTGCTGCTGACCGTTCCAGCTTCCGCTTATCAAGTGGCGACGTCCTC 243
DB |||||
DB 535 TTGGCTTTGCTGCTGCTGCTGACCGTTCCAGCTTCCGCTTATCAAGTGGCGACGTCCTC 594
QY 244 GGGATGTACATGTCAAGAACGACTGCTCCAACTCAAGCATTTGTATGAGGCGAGCGGAC 303
DB |||||
DB 595 GGGATGTACATGTCAAGAACGACTGCTCCAACTCAAGCATTTGTATGAGGCGAGCGGAC 654
QY 304 ATGATCATGCAACCCCGGGTGCCTGCTGCTGCTGGGAGAACAACTCTTCCCGCTGC 363
DB |||||
DB 655 ATGATCATGCAACCCCGGGTGCCTGCTGCTGCTGGGAGAACAACTCTTCCCGCTGC 714
QY 364 TGGGTAGCGTCAACCCCGCGCTGCGAGCTAGGAACGCGAGCGTCCCGACGACAAATA 423
DB |||||
DB 715 TGGGTAGCGTCAACCCCGCGCTGCGAGCTAGGAACGCGAGCGTCCCGACGACAAATA 774
QY 424 CGACGCCAGCTGCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 483
DB |||||
DB 775 CGACGCCAGCTGCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 834
QY 484 GACCTCTGCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 543
DB |||||
DB 835 GACCTCTGCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 894
QY 544 GAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCGCACATPACGGGTCAACGATG 603
DB |||||
DB 895 GAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCGCACATPACGGGTCAACGATG 954
QY 604 GCTTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 663
DB |||||
DB 955 GCTTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1014
QY 664 CGGATCCCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 723
DB |||||
DB 1015 CGGATCCCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1074

```

```

QY 724 CTGCGCTTACTATTCCATGCTGGGGAACCTGGGCTAAGGTTTGGATTGTGATGCTACTCTTT 783
DB |||||
DB 1075 CTGCGCTTACTATTCCATGCTGGGGAACCTGGGCTAAGGTTTGGATTGTGATGCTACTCTTT 1134
QY 784 GC 785
DB ||
DB 1135 GC 1136

RESULT 6
US-08-927-597-49
/ Sequence 49, Application US/08927597
/ Patent No. 6245503
/ GENERAL INFORMATION:
/ APPLICANT: MAERTENS, GEERT
/ APPLICANT: BOSMAN, FONS
/ APPLICANT: DE MARTYNOFF, GUY
/ APPLICANT: BUYSSE, MARIE-ANGE
/ TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
/ TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
/ NUMBER OF SEQUENCES: 111
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: NIXON & VANDERHYE P.C.
/ STREET: 1100 NORTH GLEBE ROAD
/ CITY: ARLINGTON
/ STATE: VIRGINIA
/ COUNTRY: U.S.A.
/ ZIP: 22201-4714
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25 (BPO)
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/927,597
/ FILING DATE:
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/612,973
/ FILING DATE: 11-MAR-1996
/ ATTORNEY/AGENT INFORMATION:
/ NAME: BYRNE, THOMAS E.
/ REGISTRATION NUMBER: 32,205
/ REFERENCE/DOCKET NUMBER: 1487-10
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (703) 816-4000
/ TELEFAX: (703) 816-4100
/ INFORMATION FOR SEQ ID NO: 49:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 2433 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: cDNA
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 1..2430
/ FEATURE:
/ NAME/KEY: mat_peptide
/ LOCATION: 1..2427
/
US-08-927-597-49
Query Match 98.0%; Score 778.8; DB 3; Length 2433;
Best Local Similarity 99.7%; Pred. No. 2.7e-196;
Matches 780; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 TTGGGTAAGGTCAATGATACCCCTTACATGCGGCTTCGCCGACCTCGTGGGGTACATTCGG 63
DB |||||
DB 355 TTGGGTAAGGTCAATGATACCCCTTACATGCGGCTTCGCCGACCTCGTGGGGTACATTCGG 414

```



```
QY 64 CTGCTGGGCGCCCTAGGGGGCGCTGCGAGGGCCCTGCGCGATGCGGTCCGGGTTCTG 123
DB 415 CTGCTGGGCGCCCTAGGGGGCGCTGCGAGGGCCCTGCGCGATGCGGTCCGGGTTCTG 474
QY 124 GAGGACGGGTGAATATGCAACAGGGAATTTGCCGGTGTCTCTTCTCTATCTCTCTC 183
DB 475 GAGGACGGGTGAATATGCAACAGGGAATTTGCCGGTGTCTCTTCTCTATCTCTCTC 534
QY 184 TTGGCTTTTCTGCTGCTGCTGACCGTTCCAGCTTTCGCTTATGAACTGGCCAAACGTGTCC 243
DB 535 TTGGCTTTTCTGCTGCTGCTGACCGTTTCGCTTATGAACTGGCCAAACGTGTGTCC 594
QY 244 GGGATGTACCATGTCAACGACGCTCTCCAACTCAAGCAATTTGTGTATGAGGAGCGGAC 303
DB 595 GGGATGTACCATGTCAACGACGCTCTCCAACTCAAGCAATTTGTGTATGAGGAGCGGAC 654
QY 304 ATGATCATGACACCCCGGGTGGTGGCTTGGGAGAACTCTTCCCGCTG 363
DB 655 ATGATCATGACACCCCGGGTGGTGGCTTGGGAGAACTCTTCCCGCTG 714
QY 364 TGGGTAGCGCTCACCCCGGCTCGCAGCTAGGAAACGCGAGGTCCTCCACACGCAATA 423
DB 715 TGGGTAGCGCTCACCCCGGCTCGCAGCTAGGAAACGCGAGGTCCTCCACACGCAATA 774
QY 424 CGACGCCACGTCGATTTGCTCGTTGGGGGGCTGCTTTCTGTTCCGCTATGTACGTGGG 483
DB 775 CGACGCCACGTCGATTTGCTCGTTGGGGGGCTGCTTTCTGTTCCGCTATGTACGTGGG 834
QY 484 GACCTCTGGGATCTGCTCTCTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 543
DB 835 GACCTCTGGGATCTGCTCTCTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 594
QY 544 GAGCGGTGACGACGTCGATTTGCTCAATCTATCCCGGCAACATACGAGGTCAACGATG 603
DB 895 GAGCGGTGACGACGTCGATTTGCTCAATCTATCCCGGCAACATACGAGGTCAACGATG 954
QY 604 GCTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 663
DB 955 GCTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1014
QY 664 CGGATCCACAGCTGCTGACATGTTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 723
DB 1015 CGGATCCACAGCTGCTGACATGTTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 1074
QY 724 CTGCTGCTATCTATCTGATGTTGGGAACTGGGCTTAAGCTTTGATGATGATGATGATG 783
DB 1075 CTGCTGCTATCTATCTGATGTTGGGAACTGGGCTTAAGCTTTGATGATGATGATGAT 1134
QY 784 GC 785
DB 1135 GC 1136
```

RESULT 7

```
US-09-539-601-1
; Sequence 1, Application US/09539601C
; Patent No. 6630343
; GENERAL INFORMATION:
; APPLICANT: Barten Schlager, Ralf FW
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System
; FILE REFERENCE: all sequences
; CURRENT APPLICATION NUMBER: US/09/539,601C
; CURRENT FILING DATE: 2001-08-30
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY
; EARLIER FILING DATE: 1999-04-03
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1
; LENGTH: 11076
; TYPE: DNA
; ORGANISM: Hepatitis C virus
; FEATURE:
; NAME/KEY: 5'UTR
```

```
; LOCATION: (1)..(341)
; OTHER INFORMATION: construct I389/Core-3'/wt
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (342)..(1193)
; OTHER INFORMATION: hepatitis C virus core - neomycin
; OTHER INFORMATION: phosphotransferase fusion protein
; FEATURE:
; NAME/KEY: RBS
; LOCATION: (1202)..(1812)
; OTHER INFORMATION: internal ribosome entry site from
; OTHER INFORMATION: encephalomyocarditis virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1813)..(10845)
; OTHER INFORMATION: hepatitis C virus polyprotein from core to
; OTHER INFORMATION: nonstructural protein NS5B; parental sequence
; OTHER INFORMATION: without cell culture-adaptive mutations
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: (10846)..(11076)
; US-09-539-601-1

Query Match 88.5%; Score 703.6; DB 4; Length 11076;
Best Local Similarity 93.7%; Pred. No. 3.1e-176;
Matches 733; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 4 TTGGGTAAAGTCAATGATACCCCTTACATGCGGCTTCGCGACCTCGTGGGGTACATTTCCG 63
DB 2167 TTGGGTAAAGTCAATGATACCCCTTACATGCGGCTTCGCGACCTCGTGGGGTACATTTCCG 2226
QY 64 CTGCTGGGCGCCCTTAGGGGGCGCTGCGAGGGCCCTGCGCGATGCGGCTTCG 123
DB 2227 CTGCTGGGCGCCCTTAGGGGGCGCTGCGAGGGCCCTGCGCGATGCGGCTTCG 2286
QY 124 GAGGACGGGTGAATATGCAACAGGGAATTTGCCGGTGTCTCTTCTCTATCTCTCTC 183
DB 2287 GAGGACGGGTGAATATGCAACAGGGAATTTGCCGGTGTCTCTTCTCTATCTCTCTC 2346
QY 184 TTGGCTTTTCTGCTGCTGCTGCTGACCGTTTCAGCTTTCAGCTTTCAGCTTTCAGCTTTC 243
DB 2347 TTGGCTTTTCTGCTGCTGCTGCTGACCATCCAGCTTTCAGCTTTCAGCTTTCAGCTTTC 2406
QY 244 GGGATGATCAATGTCAAGAAAGCTGCTCAACTCAAGCAATTTGTGTATGAGGAGCGGAC 303
DB 2407 GGGATGATCAATGTCAAGAAAGCTGCTCAACTCAAGCAATTTGTGTATGAGGAGCGGAC 2466
QY 304 ATGATCATGACACCCCGGGTGGTGGCTTGGGAGAACTCTTCCCGCTG 363
DB 2467 ATGATCATGACACCCCGGGTGGTGGCTTGGGAGAACTCTTCCCGCTG 2526
QY 364 TGGGTAGCGCTCACCCCGGCTGCGAGCTTAGGAAACGCGAGCTTCCCGACACGCAATA 423
DB 2527 TGGGTAGCGCTCACCCCGGCTGCGAGCTTAGGAAACGCGAGCTTCCCGACACGCAATA 2586
QY 424 CGAGCGCAAGTCAATTTGCTGTTGGGGGGCTGCTTCTGTTCCGCTATGATGAGTGGG 483
DB 2587 CGAGCGCAAGTCAATTTGCTGTTGGGGGGCTGCTTCTGTTCCGCTATGATGAGTGGG 2646
QY 484 GACCTCTGGGATCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 543
DB 2647 GATCTCTGGGATCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2706
QY 544 GAGACGGGTGAGGACCTGCAATTTGCTCAATCTATCCCGGCAACATAAGGGGTACCGGTATG 603
DB 2707 GAGACGGGTGAGGACCTGCAATTTGCTCAATCTATCCCGGCAACATAAGGGGTACCGGTATG 2766
QY 604 GCTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 663
DB 2767 GCTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2826
QY 664 CGGATCCACAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 723
```

Db 2827 CGGATCCCAAGCTGTGTGATATATGGTGGGGGGCCCAATTTGGGAGTCTTAGCGGGC 2886
QY 724 CTGCGCTACTATTCCATGGTGGGAACTGGGCTAAGGTTTGAATTGTGATGCTACTCTTT 783
Db 2887 CTTGGCTACTATTCCATGGTGGGAACTGGGCTAAGGTTTGTGATTTGTGATGCTACTCTTT 2946
QY 784 GC 785
Db 2947 GC 2948
RESULT 8
US-09-539-601-19
; Sequence 19, Application US/09539601C
; Patent No. 6630343
; GENERAL INFORMATION:
; APPLICANT: Bartenschlager, Ralf FW
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System
; FILE REFERENCE: all sequences
; CURRENT APPLICATION NUMBER: US/09/539,601C
; CURRENT FILING DATE: 2001-08-30
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY
; EARLIER FILING DATE: 1999-04-03
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 11076
; TYPE: DNA
; ORGANISM: Hepatitis C virus
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (1)..(341)
; OTHER INFORMATION: construct I389/Core-3'/9-13P
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (342)..(1193)
; OTHER INFORMATION: hepatitis C virus core - neomycin
; OTHER INFORMATION: phosphotransferase fusion protein
; FEATURE:
; NAME/KEY: RBS
; LOCATION: (1202)..(1812)
; OTHER INFORMATION: internal ribosome entry site from
; OTHER INFORMATION: encephalomyocarditis virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1813)..(10845)
; OTHER INFORMATION: hepatitis C virus polyprotein from core to
; OTHER INFORMATION: nonstructural protein NS5B; carries cell
; OTHER INFORMATION: culture-adaptive mutations from clone 9-13P
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: (10846)..(11076)
US-09-539-601-19
Query Match 88.5%; Score 703.6; DB 4; Length 11076;
Best Local Similarity 93.7%; Pred. No. 3.1e-176;
Matches 733; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
QY 4 TTGGGTAAGGTCTATCGATACCTTTACATGGGGCTTCGCGGCTTCGCGGCTACATTTCCG 63
Db 2167 TTGGGTAAGGTCTATCGATACCTTCACGTCGGCTTCGCGGCTACATTTCCG 2226
QY 64 CTGCTGGCGCCCGCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGGTTCTG 123
Db 2227 CTGCTGGCGCCCGCTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGGTTCTG 2286
QY 124 GAGACGGCGTGAACATATGCAACGGGAATTTGCCGGTTCCTTTCTATCTTCTC 183
Db 2287 GAGACGGCGTGAACATATGCAACGGGAATTTGCCGGTTCCTTTCTATCTTCTC 2346
QY 184 TTGGCTTGTCTGCTCTGACCGTTCCAGCTTCCAGCTTATGAAGTGGCAAGCTGCC 243
Db 2347 TTGGCTTGTCTGCTCTGACCGTTCCAGCTTATGAAGTGGCAAGCTATCC 2406

QY 244 GGGATGTACCATGTCAAGACAGCTGTCTCCAACTCAAGCATTTGTGTATGAGCAGCGAC 303
Db 2407 GGGATGTACCATGTCAAGACAGCTGTCTCCAACTCAAGCATTTGTGTATGAGCAGCGAC 2466
QY 304 ATGATCATGACACACCCCGGGTGGTCCCTGGCTTGGGAGAACAACTTCCCGCTGC 363
Db 2467 ATGATCATGACACACCCCGGGTGGTCCCTGGCTTGGGAGAACAACTTCCCGCTGC 2526
QY 364 TGGGTAGCGCTCACCCCGACGCTCGACGTAGGAAACGACGCTCCCAACACGACAAATA 423
Db 2527 TGGGTAGCGCTCACCTCCACGCTCGGACAGGACGCTAGCGTCCCACTACGACGATA 2586
QY 424 CGACGCCACGCTCGAATTTGCTCGTTGGGGCGGCTGCTTTCTGTTCCGCTATGTACGTGGGG 483
Db 2587 CGACGCCATGTCTGATTTGCTCGTTGGGGCGGCTGCTTTCTGTTCCGCTATGTACGTGGGG 2646
QY 484 GACCTCTGGGATCTGCTTCTCTGCTCCAGCTGTTTACCATCTCGCTCGCGCGGAT 543
Db 2647 GATCTCTGGGATCTGTTTCTCTGCTCGCCAGCTGTTTCACTTCTCGCTCGCGCGAC 2706
QY 544 GAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAAACGGTCAACGATG 603
Db 2707 GAGACGGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCCACATAAACGGTCAACGATG 2766
QY 604 GCTTGGGATATGATGAACTGCTCGCTTACCAACGGGCCCTGGTGGTATGGCAGCTGCTC 663
Db 2767 GCTTGGGATATGATGAACTGCTCGCTTACCAACGGGCCCTGGTGGTATGGCAGTACTC 2826
QY 664 CGGATCCCAAGCTGTCTGGACATGGTGGGGGGCCCATTTGGGGAGTCTCTGGCGGT 723
Db 2827 CGGATCCCAAGCTGTCTGGATATGGTGGGGGGCCCATTTGGGGAGTCTCTAGCGGGC 2886
QY 724 CTGCGCTACTATTCCATGGTGGGAACTGGGCTAAGGTTTGAATTCATGCTACTCTTT 783
Db 2887 CTTGCGCTACTATTCCATGGTGGGAACTGGGCTAAGGTTTGAATTCATGCTACTCTTT 2946
QY 784 GC 785
Db 2947 GC 2948

RESULT 9

US-09-539-601-25
; Sequence 25, Application US/09539601C
; Patent No. 6630343
; GENERAL INFORMATION:
; APPLICANT: Bartenschlager, Ralf FW
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System
; FILE REFERENCE: all sequences
; CURRENT APPLICATION NUMBER: US/09/539,601C
; CURRENT FILING DATE: 2001-08-30
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY
; EARLIER FILING DATE: 1999-04-03
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 25
; LENGTH: 11076
; TYPE: DNA
; ORGANISM: Hepatitis C virus
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (1)..(341)
; OTHER INFORMATION: construct I389/Core-3'/5.1
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (342)..(1193)
; OTHER INFORMATION: hepatitis C virus core - neomycin
; OTHER INFORMATION: phosphotransferase fusion protein
; FEATURE:
; NAME/KEY: RBS
; LOCATION: (1202)..(1812)
; OTHER INFORMATION: internal ribosome entry site from

```

; OTHER INFORMATION: encephalomyocarditis virus
; NAME/KEY: CDS
; LOCATION: (1813)..(10845)
; OTHER INFORMATION: hepatitis C virus polypeptide from core to
; OTHER INFORMATION: nonstructural protein NS5B; carries cell
; OTHER INFORMATION: culture-adaptive mutations of clone 5.1
; FEATURE:
; NAME/KEY: 3' UTR
; LOCATION: (10846)..(11076)
US-09-539-601-25

Query Match      88.5%; Score 703.6; DB 4; Length 11076;
Best Local Similarity 93.7%; Pred. No. 3.1e-176;
Matches 733; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

Qy 4 TTGGGTAAGGTCATCGATACCTTACATCGCGGCTTGGCGGACCTCGTGGGTACATTCGG 63
Db 2167 TTGGGTAAGGTCATCGATACCTTACATCGCGGCTTGGCGGACCTCGTGGGTACATTCGG 2226

Qy 64 CTGCTGCGGCGCCCTTAGGGGGCGCTGCCAGGGCCCTGGGCGCATGGCGTTCGGGTTCG 123
Db 2227 CTGCTGCGGCGCCCTTAGGGGGCGCTGCCAGGGCCCTGGGCGCATGGCGTTCGGGTTCG 2286

Qy 124 GAGGACGCGGTGAACATGCAACAGGGAATTTGCCCGGTGCTCTTTCTATCTTCCTC 183
Db 2287 GAGGACGCGGTGAACATGCAACAGGGAATTTGCCCGGTGCTCTTTCTATCTTCCTC 2346

Qy 184 TTGGCTTTGCTGCTGTCAGCTTCCAGCTTCCAACTGCTTGAAGTGGCGGACGTGTC 243
Db 2347 TTGGCTTTGCTGCTGTCAGCTTCCAGCTTCCAACTGCTTGAAGTGGCGGACGTGTC 2406

Qy 244 GGGATGTACCATGTCACCAAGACGCTTCCAACTGCTTGAAGTGGCGGACGTGTC 303
Db 2407 GGGATGTACCATGTCACCAAGACGCTTCCAACTGCTTGAAGTGGCGGACGTGTC 2466

Qy 304 ATGATCATGCAACACCCCGGGTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 363
Db 2467 ATGATCATGCAACACCCCGGGTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2526

Qy 364 TGGGTAGGCTTACCCCGGCTGCGAGTGGAGGACGCGTCCCAACGACGACATA 423
Db 2527 TGGGTAGGCTTACCCCGGCTGCGAGTGGAGGACGCGTCCCAACGACGACATA 2586

Qy 424 CGAGCCACGTCGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 483
Db 2587 CGAGCCACGTCGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2646

Qy 484 GACCTCTGCGGATCTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 543
Db 2647 GACCTCTGCGGATCTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2706

Qy 544 GAGAGGTGCGAGGCTGCAATGCTCAATCTATCCCGGCGACATTAACGGGTACCGGTATG 603
Db 2707 GAGAGGTGCGAGGCTGCAATGCTCAATCTATCCCGGCGACATTAACGGGTACCGGTATG 2766

Qy 604 GCTTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 663
Db 2767 GCTTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2826

Qy 664 CCGATCCCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 723
Db 2827 CCGATCCCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2886

Qy 724 CTCGCTACTATTTCCATGGTGGGAACTGGGGTAAAGTTTGAATTTGAATGATGATCTCTTT 783
Db 2887 CTCGCTACTATTTCCATGGTGGGAACTGGGGTAAAGTTTGAATTTGAATGATGATCTCTTT 2946

Qy 784 GC 785
Db 2947 GC 2948

```

```

RESULT 10
US-09-539-601-31
; Sequence 31, Application US/09539601C
; Patent No. 6630343
; GENERAL INFORMATION:
; APPLICANT: Bartenschlager, Ralf FW
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System
; FILE REFERENCE: all sequences
; CURRENT APPLICATION NUMBER: US/09/539,601C
; EARLIER FILING DATE: 2001-08-30
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY
; EARLIER FILING DATE: 1999-04-03
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 31
; LENGTH: 11076
; TYPE: DNA
; ORGANISM: Hepatitis C virus
; FEATURE:
; NAME/KEY: 5' UTR
; LOCATION: (1)..(341)
; OTHER INFORMATION: construct 1389/Core-3'/19
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (342)..(1193)
; OTHER INFORMATION: hepatitis C virus core - neomycin
; OTHER INFORMATION: phosphotransferase fusion protein
; FEATURE:
; NAME/KEY: RBS
; LOCATION: (1202)..(1812)
; OTHER INFORMATION: internal ribosome entry site from
; OTHER INFORMATION: encephalomyocarditis virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1813)..(10845)
; OTHER INFORMATION: hepatitis C virus polypeptide from core to
; OTHER INFORMATION: nonstructural protein NS5B; carries cell culture
; OTHER INFORMATION: adaptive mutations from clone no. 19
; NAME/KEY: 3' UTR
; LOCATION: (10846)..(11076)
US-09-539-601-31

Query Match      88.5%; Score 703.6; DB 4; Length 11076;
Best Local Similarity 93.7%; Pred. No. 3.1e-176;
Matches 733; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

Qy 4 TTGGGTAAGGTCATCGATACCTTACATCGCGGCTTGGCGGACCTCGTGGGTACATTCGG 63
Db 2167 TTGGGTAAGGTCATCGATACCTTACATCGCGGCTTGGCGGACCTCGTGGGTACATTCGG 2226

Qy 64 CTGCTGCGGCGCCCTTAGGGGGCGCTGCCAGGGCCCTGGGCGCATGGCGTTCGGGTTCG 123
Db 2227 CTGCTGCGGCGCCCTTAGGGGGCGCTGCCAGGGCCCTGGGCGCATGGCGTTCGGGTTCG 2286

Qy 124 GAGGACGCGGTGAACATGCAACAGGGAATTTGCCCGGTGCTCTTTCTATCTTCCTC 183
Db 2287 GAGGACGCGGTGAACATGCAACAGGGAATTTGCCCGGTGCTCTTTCTATCTTCCTC 2346

Qy 184 TTGGCTTTGCTGCTGTCAGCTTCCAGCTTCCAACTGCTTGAAGTGGCGGACGTGTC 243
Db 2347 TTGGCTTTGCTGCTGTCAGCTTCCAGCTTCCAACTGCTTGAAGTGGCGGACGTGTC 2406

Qy 244 GGGATGTACCATGTCACCAAGACGCTTCCAACTGCTTGAAGTGGCGGACGTGTC 303
Db 2407 GGGATGTACCATGTCACCAAGACGCTTCCAACTGCTTGAAGTGGCGGACGTGTC 2466

Qy 304 ATGATCATGCAACACCCCGGGTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 363
Db 2467 ATGATCATGCAACACCCCGGGTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2526

Qy 364 TGGGTAGGCTTACCCCGGCTGCGAGTGGAGGACGCGTCCCAACGACGACATA 423

```

Db	2527	TGGGTAGCGTCACTCCACAGCTCGCGGCCAGGAAACGTAGCTGCCCACTACAGCAGATA	2586
Qy	424	CGACGCCACGTCGATTTTGCTGTTGGGGCGGCTCTTTCTGTTCCGCTATGTATACGTGGG	483
Db	2587	CGACGCCATGTGCGATTTGCTGTTGGGGCGGCTCTCTCTGCTCCGCTATGTATACGTGGGA	2646
Qy	484	GACCTTCGCGGATCTGTCTTCCTCGTCTGCCAGCTGTTTCCACATCTCGGCTCGCCGCGCAT	543
Db	2647	GATCTCTGCGGATCTGTTTCTCGTGGCCAGCTGTTTCACTTCTCGCTTCGCGCGGAC	2706
Qy	544	GAGACGGTGCAGGACTGCAATTGCTCAATCTATATCCCGGCCACATAACGGGTCACCGTATG	603
Db	2707	GAGACAGTACAGGACTGCAATTGCTCAATATATCCCGGCCACAGTTCACCGTATG	2766
Qy	604	GCTTGGGATATGATGATGAACCTGTCGCTACAAACGGCCCTGTTGGTATCGACGTCGTC	663
Db	2767	GCTTGGGATATGATGATGAACCTGTCGCTACCACTGACAGAGCCCTAGTGGTATCGCAGTTACTC	2826
Qy	664	CGGATCCCAAGCTGTCGTGGACATCGTGGCGGGGCCCAATTGGGGAGTTCCTGGCGGGT	723
Db	2827	CGGATCCCAAGCTGTCGTGGATATGTTGGCGGGGCCCAATTGGGGAGTTCCTAGCGGGC	2886
Qy	724	CTCGCCTACTATTCCATGGTGGGAACTGGGCTAAGGTTTTTGATTTGATGTGCTACTCTTTT	783
Db	2887	CTTGGCTACTATTCCATGGTGGGAACTGGGCTAAGGTTCTGATTTGATGTGCTACTCTTTT	2946
Qy	784	GC 785	
Db	2947	GC 2948	

RESULT 11

US-08-470-426B-17
; Sequence 17, Application US/08470426B
; Patent No. 5856458
; GENERAL INFORMATION:
; APPLICANT: Okamoto, Hiroaki
; APPLICANT: Nakamura, Tetsuo
; TITLE OF INVENTION: OLIGONUCLEOTIDE PRIMERS, AND THEIR
; TITLE OF INVENTION: APPLICATION FOR HIGH-FIDELITY DETECTION OF NON-A, NON-B
; TITLE OF INVENTION: HEPATITIS VIRUS
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Beveridge, DeGrandi, Weilacher & Young,
; ADDRESSEE: L.L.P.
; STREET: 1850 M Street, N.W., Suite 800
; CITY: Washington
; STATE: DC
; COUNTRY: USA

```

? STRANDEDNESS: single
? TOPOLOGY: unknown
? MOLECULE TYPE: DNA (genomic)
US-08-470-4268-17

```

Query Match	87.9%	Score 698.8;	DB 2;	Length 1539;
Best Local Similarity	93.4%	Pred. No. 3.1e-175;		
Matches 730; Conservative	0;	Mismatches 52;	Indels 0;	Gaps 0;

Qy	4	TTGGGTAAGGTCAATCGATACCCCTTACATGCGGGCTTGGCCGACCTGTCGGGGTACATTCGG	63
Db	355	TTGGGTAAGGTCAATCGATACCCCTTACATGCGGGCTTGGCCGATCTCATGGGGTATATTTCCC	414
Qy	64	CTCGTCCGGCCGCCCTTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGCTCCGGGTCTCG	123
Db	415	CTCGTCCGGCCGCCCTTAGGGGGCGCTGCCAGGGCCTTGGCAACGGGTGTCCGGGTCTCG	474
Qy	124	GAGGACGGCGTGAACATATGCAACAGGGAAATTTGGCCGGGTGCTCTTTCTCTATCTTCCTC	183
Db	475	GAGGACGGCGTGAACATATGCAACAGGGAACTTTGCCCGGTGCTCTTTCTCTATCTTCCTC	534
Qy	184	TTGGCTTTGCTGCTGCTGCTGACGGTTCAGCTTCGGCTTATGCAAGTGGCGCAAGGTGCC	243
Db	535	TTGGCTTTGCTGCTGCTGCTGACCATCCGAGCTTCGGCTTATGAGTGGCGCAACGTGTGCC	594
Qy	244	GGGATGTACCAATGTCAAGAACGACTGCTCCAACTCAAGCAATTGTGTATGAGGCAGACGGAC	303
Db	595	GGGATATACCAATGTCAAGAACGACTGCTCCAACTCAAGCAATTGTGTATGAGGCAGACGGAC	654
Qy	304	ATGATATCATGACACCCCGGGTGGGTGCCCTGCTGGTTCGGGAGAACAACTCTTCCCGGTGCC	363
Db	655	ATGATATCATGATACTCCCGGGTGGGTGCCCTGCTGGTTCGGGAGGACAAACAGTCCCGGTGCC	714
Qy	364	TGGGTAGCGCTCACCCGCCACGCTCGCAGCTTAGGAACGCCAGCGTCCGCCACACGACAATA	423
Db	715	TGGGTAGCGCTCACTCCACGCTCGCGCCAGGAATGCCAGGGTCCCCCTACTGACACHATA	774
Qy	424	CGACGCCACGTCGAATTTGCTCGTTGGGGGGGCTGCTTTCTGTTCCGCTATGTATAGTGGGG	483
Db	775	CGACGCCACGTCGAATTTGCTCGTTGGGGGGGCTGCTTTCTGCTCGCTATGTATAGTGGGG	834
Qy	484	GACTCTCGGGATCTGTCTCTCTCGTCTCCAGCTGTTCACCACTCTCGCTTCGCCCGGCAT	543
Db	835	GATCTCTCGGGATCTGTCTCTCTCGTCTCCAGCTGTTCACCTCTCTCGCTTCGCCCGGCAT	894
Qy	544	GAGACGGTCAGGACTGCAATTGCTCAATCTATTCGCCGCCACATAACGGGTCAACCGTATG	603
Db	895	GAGACAGTCAGGACTGCAACTGCTCAATCTATTCGCCGCCATTTATCAGGTCAACCGCATG	954
Qy	604	GCTTGGGATATGATGATGAACCTGGTCGCCCTACAAACGGCCCTGGTGGTATACGACGTGCTC	663
Db	955	GCTTGGGATATGATGATGAACCTGGTCGCCCTACAAACGGCCCTAGTGGTGTGCGAGTTGCTC	1014
Qy	664	CGGATCCACACAGCTGTGTTGGACATGGTGGCGGGGGGCCCATTTGGGGAGTCTCTGGGGGGT	723
Db	1015	CGGATCCACACAGCTGTGTTGGACATGGTGGCGGGGGGCCCATCTGGGGAGTCTCTGGGGGGC	1074
Qy	724	CTCGCTACTATTTCATGTTGGGGAACCTGGGCTTAAGGTTTGTGATGTGATGCTACTCTTT	783
Db	1075	CTTGGCTACTATTTCATGTTAGGGAACCTGGGCTTAAGGTTCTGATGTGTGGCGTACTCTTTC	1134
Qy	784	GC 785	
Db	1135	GC 1136	

RESULT 12

US-08-470-426B-14
; Sequence 14, Application US/08470426B
; Patent No. 5856458
; GENERAL INFORMATION:
; APPLICANT: Okamoto, Hiroaki
; APPLICANT: Nakamura, Tetsuo

;; TITLE OF INVENTION: OLIGONUCLEOTIDE PRIMERS, AND THEIR
;; TITLE OF INVENTION: APPLICATION FOR HIGH-FIDELITY DETECTION OF NON-A, NON-B
;; TITLE OF INVENTION: HEPATITIS VIRUS
;; NUMBER OF SEQUENCES: 33
;; CORRESPONDENCE ADDRESSES:
;; ADDRESS: Beveridge, DeGrandi, Weillacher & Young,
;; ADDRESS: L.L.P.
;; STREET: 1850 M Street, N.W., Suite 800
;; CITY: Washington
;; STATE: DC
;; COUNTRY: USA
;; ZIP: 20036
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent In Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; FILING DATE: 06-JUN-1995
;; PRIORITY APPLICATION NUMBER: US/08/470,426B
;; CLASSIFICATION: 536
;; PRIOR APPLICATION NUMBER:
;; FILING DATE: 12-JUN-1990
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Weillacher, Robert G.
;; REGISTRATION NUMBER: 20,531
;; REFERENCE/DOCKET NUMBER: 06/59-47083.1
;; TELEPHONE: (202) 659-2811
;; TELEFAX: (202) 659-1462
;; INFORMATION FOR SEQ ID NO: 14:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1863 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: unknown
;; MOLECULE TYPE: DNA (genomic)
US-08-470-426B-14

Query Match 87.9%; Score 698.8; DB 2; Length 1863;
Best Local Similarity 93.4%; Pred. No. 3.3e-175;
Matches 730; Conservative 0; Mismatches 52; Indels 0; Gaps 0;
QY 4 TTGGTAAGGTGATGATACCTTACATGCGGCTTCCGCGACTCGTGGGGTACATTCGG 63
DB 679 TTGGTAAGGTGATGATACCTTACATGCGGCTTCCGCGACTCGTGGGGTACATTCGG 738
QY 64 CTGCTCGGCGCCCCCTAGGGGCGCTGCGAGGGCCCTGGCGCATGGCGTCCGGGTTCTG 123
DB 739 CTGCTCGGCGCCCCCTAGGGGCGCTGCGAGGGCCCTGGCGCATGGCGTCCGGGTTCTG 798
QY 124 GAGGACGGGTGAACTATGTCACAGGGAATTTCCCGGTTGCTCTTCTATCTCTCTC 183
DB 799 GAGGACGGGTGAACTATGTCACAGGGAATTTCCCGGTTGCTCTTCTATCTCTCTC 858
QY 184 TTGGCTTTGCTCTCTCTGCTGACCGTTCCAGCTTCCAGCTTATGAGTGGCAACGTGTC 243
DB 859 TTGGCTTTGCTCTCTCTGTTGACCATCCAGCTTCCAGCTTATGAGTGGCAACGTGTC 918
QY 244 GGGATGTACCATGTGTCACGAAACGCTCTCAACTCAAGCAATTTGTATGAGGACGGGAC 303
DB 919 GGGATGTACCATGTGTCACGAAACGCTCTCAACTCAAGCAATTTGTATGAGGACGGGAC 978
QY 304 ATGATCATGCACACCCCGGGTGGTCCCTGGGTTGGGAGNACAACTCTTCCCGCTGC 363
DB 979 ATGATCATGCACACTCCCGGGTGGTCCCTGGGTTGGGAGNACAACTCCCGCTGC 1038
QY 364 TGGGTAGCGCTCACCCCGGCTCGGAGCTAGGAAACGCGAGCTCCCGGAGCAATA 423
DB 1039 TGGGTAGCGCTCACCTCCCGGCTCGGAGCTAGGAAACGCGAGCTCCCGGAGCAATA 1098
QY 424 CGAAGCCACGTGCAATTTGCTGTTGGGGCGGCTGCTTTCTGTTCCGCTATGTAAGTGGGG 483

DB 1099 CGAAGCCACGTGCAATTTGCTGTTGGGGCGGCTGCTTTCTGCTCGCTATGTAAGTGGGG 1158
QY 484 GACCTCTGGGATCTGCTCTTCTGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCG 543
DB 1159 GATCTCTGGGATCTGCTCTTCTGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCG 1218
QY 544 GAGAGCGGTGCGAGGACTGCAATTTGCTCAATCTATCCCGGCCACATACCGGGTCAACGGTATG 603
DB 1219 GAGACAGTGCAGGACTGCAATTTGCTCAATCTATCCCGGCCATTTATCAGGTCAACGGATG 1278
QY 604 GCTTGGGATATGATGATGAACTGGTGGCTTACAAAGCGGCTCGTGGTATGCGAGCTGCTC 663
DB 1279 GCTTGGGATATGATGATGAACTGGTGGCTTACAAAGCGGCTCGTGGTATGCGAGTTCCTC 1338
QY 664 CGGATCCCAAGCTGCTGTCGACATGTCGCGGGGCGGCTTGGGGAGTCTTGGCGGGT 723
DB 1339 CGGATCCCAAGCTGCTGTCGACATGTCGCGGGGCGGCTTGGGGAGTCTTGGCGGGC 1398
QY 724 CTCGCTACTATTTCCATGCTGGGGAACCTGGGCTTAAGGTTTTGATGTTGATGCTCTCTT 783
DB 1399 CTTCGCTACTATTTCCATGCTGGGGAACCTGGGCTTAAGGTTTTGATGTTGATGCTCTCTT 1458
QY 784 GC 785
DB 1459 GC 1460
RESULT 13
US-08-191-160-21
; Sequence 21, Application US/08191160
; Patent No. 6210675
; GENERAL INFORMATION:
; APPLICANT: Highfield, Peter Edmund
; APPLICANT: Rodgers, Brian Colin
; APPLICANT: Tedder, Richard Seton
; APPLICANT: Barbara, John Anthony James
; TITLE OF INVENTION: Viral Agent
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rothwell, Figg, Ernst & Kurz
; STREET: 1700 K Street
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy diskette, 5.25 inch, 360 Kb storage
; COMPUTER: IBM AT compatible
; OPERATING SYSTEM: MS-DOS V3.2
; SOFTWARE: Wordperfect 5.0 (DOS text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/191,160
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/628,516
; FILING DATE: 17 DEC 1990
; APPLICATION NUMBER: UK 89 28 562.1
; FILING DATE: 18 DEC 1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: UK 90 04 414.0
; FILING DATE: 27 FEB 1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: UK 90 04 814.1
; FILING DATE: 03 MAR 1990
; ATTORNEY/AGENT INFORMATION:
; NAME: E. Anthony Figg
; REGISTRATION NUMBER: 27,195
; REFERENCE/DOCKET NUMBER: 1645-103A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 833-5740
; TELEFAX: (202) 833-5744

INFORMATION FOR SEQ ID NO: 21:

SEQUENCE CHARACTERISTICS:
LENGTH: 2116 base pairs
TYPE: nucleotide with corresponding protein
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to genomic RNA
ORIGINAL SOURCE:
ORGANISM: human; serum infectious for PT-NANBH
IMMEDIATE SOURCE:
LIBRARY: cDNA clones from 5' end of the genome
FEATURE:
LOCATION: from 308 to 2116 bp start of the PT-NANBH
LOCATION: polyprotein
OTHER INFORMATION: viral structural and non-structural
OTHER INFORMATION: proteins
US-08-191-160-21

Query Match 87.7%; Score 697.2; DB 3; Length 2116;
Best Local Similarity 93.2%; Pred. No. 9e-175;
Matches 729; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY	4	TTGGGTAAGTGCATCGATACCCCTTACATGGCGCTTCGCGACCTCGTGGGTACATTCOG	63
DB	662	TTGGGTAAGTGCATCGATACCCCTTACATGGCGCTTCGCGACCTCGTGGGTACATTCOG	721
QY	64	CTCGTCGGGGCCCCCTTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTTCTG	123
DB	722	CTCGTCGGGGCCCCCTTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTTCTG	781
QY	124	GAGGACGGGTGAACATATGACACAGGGAATTTGCCGGTTGCTTTCTATCTTCCTC	183
DB	782	GAGGACGGGTGAACATATGACACAGGGAATTTGCCGGTTGCTTTCTATCTTCCTC	841
QY	184	TTGGCTTTGCTGTCCTGTCGACGCTTCAGCTTCGCTTATCAAGTGGCGACGTTCC	243
DB	842	TTGGCTTTGCTGTCCTGTCGACGCTTCAGCTTCGCTTATCAAGTGGCGACGTTCC	901
QY	244	GGGATGATACATGTCACAGAACGCTCTCCAACTCAAGCATCTGTATGAGGACGGGAC	303
DB	902	GGGATGATACATGTCACAGAACGCTCTCCAACTCAAGCATCTGTATGAGGACGGGAC	961
QY	304	ATGATCATGCACACCCCGGGTGGCTGCTTGGGAGAACAACTCTTCCCGCTGC	363
DB	962	ATGATCATGCACACCCCGGGTGGCTGCTTGGGAGAACAACTCTTCCCGCTGC	1021
QY	364	TGGGTAGCCCTACCCCGACGCTCGGAGTAGAAGACCGGCGTCCCGACAGACAATA	423
DB	1022	TGGGTAGCCCTACCCCGACGCTCGGAGTAGAAGACCGGCGTCCCGACAGACAATA	1081
QY	424	CGAGCGCAGTTCGATTTGCTCGTTGGGGGGGCTTCTTCTGTTCCGCTATGTACGTGGG	483
DB	1082	CGAGCGCAGTTCGATTTGCTCGTTGGGGGGGCTTCTTCTGTTCCGCTATGTACGTGGG	1141
QY	484	GACCTCTGGGATCTCTTCTCTCCAGCTGTTCACCACTCTCGCCTCGCGGCAAT	543
DB	1142	GATCTCTGGGATCTCTTCTCTCCAGCTGTTCACCACTCTCGCCTCGCGGCAAT	1201
QY	544	GAGACGCTGACGAGTGCATTTGCTCAATCTATCCCGGCGACATAACGGGTGACCGTATG	603
DB	1202	CAGACGCTGACGAGTGCATTTGCTCAATCTATCCCGGCGACATAACGGGTGACCGTATG	1261
QY	604	GCTTGGGATATGATGATGAACCTGCTGCTTACAAACCGGCGCTGGTGTATCGGAGCTCTC	663
DB	1262	GCTTGGGATATGATGATGAACCTGCTGCTTACAAACCGGCGCTGGTGTATCGGAGCTCTC	1321
QY	664	CGGATCCCAAGCTGCTGTGGACATGGTGGGGGGGCCCATTTGGGGAGTCTTGGCGGGT	723
DB	1322	CGGATCCCAAGCTGCTGTGGACATGGTGGGGGGGCCCATTTGGGGAGTCTTGGCGGGC	1381
QY	724	CTCGCTACTATTTCCATGTGGGGAATCTGGGCTTAAGGTTTGTATGTGATGCTACTTT	783
DB	1382	CTTGCCTACTATTTCCATGTGGGGAATCTGGGCTTAAGGTTTGTATGTGATGCTACTTT	1441

QY 784 GC 785
DB 1442 GC 1443

RESULT 14

US-08-081-072-15
Sequence 15, Application US/08081072
Patent No. 5641654
GENERAL INFORMATION:
APPLICANT: No. 5641654oru MAKI, Kenjiro YAMAGUCHI, Ayumi
TITLE OF INVENTION: NON-A NON-B HEPATITIS-SPECIFIC
TITLE OF INVENTION: ANTIGEN AND ITS USE IN HEPATITIS
TITLE OF INVENTION: DIAGNOSIS
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dike, Bronstein, Roberts & Cushman
STREET: 130 Water Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-4280
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk, 3.50inch, 1.4Mb storage
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh Classic
SOFTWARE: Microsoft Word Version 4.0A
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/081,072
FILING DATE: June 22, 1993
CLASSIFICATION: 435
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 932 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to genomic RNA
US-08-081-072-15

Query Match 87.5%; Score 695.6; DB 1; Length 932;
Best Local Similarity 93.1%; Pred. No. 1.8e-174;
Matches 728; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY	4	TTGGGTAAGTGCATCGATACCCCTTACATGGCGCTTCGCGACCTCGTGGGTACATTCOG	63
DB	9	TTGGGTAAGTGCATCGATACCCCTTACATGGCGCTTCGCGACCTCGTGGGTACATTCOG	68
QY	64	CTCGTCGGGGCCCCCTTAGGGGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTTCTG	123
DB	69	CTTGTGCGGGCCCCCTTAGGGGGTGTCTGCCAGGGCCCTGGCACATGGTGTCCGGGTTCTG	128
QY	124	GAGGACGGGTGAACATATGACACAGGGAATTTGCCGGTTGCTTTCTATCTTCCTC	183
DB	129	GAGGACGGGTGAACATATGACACAGGGAATTTGCCGGTTGCTTTCTATCTTCCTC	188
QY	184	TTGGCTTTGCTGTCCTGTCGACGCTTCAGCTTCGCTTATGAAAGTGGCGACGTTCC	243
DB	189	TTGGCTTTGCTGTCCTGTCGACGCTTCAGCTTCGCTTATGAAAGTGGCGACGTTCC	248
QY	244	GGGATGATACATGTCACAGAACGCTCTCCAACTCAAGCATCTGTATGAGGACGGGAC	303
DB	249	GGGATGATACATGTCACAGAACGCTCTCCAACTCAAGCATCTGTATGAGGACGGGAC	308
QY	304	ATGATCATGCACACCCCGGGTGGCTGCTTCCAGTCTGGGAGAACAACTCTTCCCGCTGC	363
DB	309	ATGATCATGCATACCCCGGGTGGCTGCTTCCAGTCTGGGAGAACAACTCTTCCCGCTGC	368
QY	364	TGGGTAGCCCTACCCCGACGCTCGGAGTAGAAGACCGGCGTCCCGACAGACAATA	423
DB	369	TGGGTAGCCCTACCCCGACGCTTAGCGGCGGAGAACACCGGCGTCCCGACAGACAATA	428

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2004, 08:32:08 ; Search time 1473.55 Seconds
(without alignments)
14206.949 Million cell updates/sec

Title: US-09-899-303a-9
Perfect score: 483
Sequence: 1 ATGCCCGGTTGCTCTTCTC.....TGATGAACCTGGTCTAATAG 483

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_stb.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_stb.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rod.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	483	100.0	483	6	A48671
2	483	100.0	483	6	AR157327 Sequence 9
3	483	100.0	483	6	AX452758 Sequence
4	483	100.0	483	6	AX685010 Sequence
5	483	100.0	483	6	AX794852 Sequence
6	474.2	98.2	480	6	A48673 Sequence 11
7	474.2	98.2	480	6	AR157328 Sequence
8	474.2	98.2	480	6	AX452760 Sequence
9	474.2	98.2	480	6	AX685012 Sequence
10	474.2	98.2	480	6	AX794854 Sequence
11	438.2	90.7	1880	14	HPC5TRJ4
12	435	90.1	1880	14	HPC5TRJ4
13	435	90.1	2540	6	E04260
14	435	90.1	2540	6	E04805
15	435	90.1	2540	6	E07391
16	435	90.1	9448	14	HPCJ483
17	433.4	89.7	1539	6	AR027786
18	433.4	89.7	1863	6	AR027783
19	433	89.6	9580	14	AF054250
20	431.8	89.4	8780	14	AF054257
21	431.8	89.4	8780	14	AF054258
22	430.2	89.1	8780	14	AF054255
23	430.2	89.1	8781	14	AF054254
24	430.2	89.1	9379	14	AF165052
25	430.2	89.1	9460	14	HPCJ491
26	430.2	89.1	9595	6	AR119832
27	430.2	89.1	9595	14	AF054247
28	430.2	89.1	9596	14	AF054249
29	430.2	89.1	9599	6	AR119833
30	428.6	88.7	577	6	E04085
31	428.6	88.7	618	14	HPCB1P12
32	428.6	88.7	8779	14	AF054251
33	428.6	88.7	8780	14	AF054252
34	428.6	88.7	9377	14	AF207756
35	428.6	88.7	9410	14	HPCB1R2
36	428.6	88.7	9435	14	AB049093
37	428.6	88.7	9595	14	AF054248
38	427	88.4	1595	14	HPCNS1SPF
39	427	88.4	1635	14	HPCNS1SPD
40	427	88.4	3296	14	AB008446
41	427	88.4	8780	14	AF054259
42	427	88.4	8781	14	AF054256
43	427	88.4	9361	14	AF483269
44	427	88.4	9369	14	AF165054
45	427	88.4	9379	14	AF165051

ALIGNMENTS

RESULT 1					
A48671	A48671	Sequence 9 from Patent WO9604385.	483 bp	DNA	linear
LOCUS					
DEFINITION	A48671				
ACCESSION	A48671				
VERSION	A48671.1	GI:2302384			
KEYWORDS					
SOURCE		unidentified			
ORGANISM		unidentified			
REFERENCE		1 (bases 1 to 483)			
AUTHORS		Maertens,G., Bosman,F., De,M.G. and Buyse,M.			
TITLE		PURIFIED HEPATITIS C VIRUS ENVELOPE PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE			
JOURNAL		Patent: WO 9604385-A 9 15-FEB-1996;			

Db 1207 AGCTGTTTACCTTCTCGCCTCGCGCATGAGACAGTGCAGGACTGCCTCAATCT 1266
Qy 422 ATCCCGGCCATGTATCAGGTCAACGCGATGGCTTGGGATATGATGATGAACTGGTC 476
Db 1267 ATCCCGGCCATTTATCAGGTCAACGCGATGGCTTGGGATATGATGATGAACTGGTC 1321

Search completed: June 16, 2004, 11:39:17
Job time : 1474.55 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2004, 07:39:53 ; Search time 181.729 Seconds
(without alignments)
11290.892 Million cell updates/sec

Title: US-09-899-303A-9

Perfect score: 483

Sequence: 1 ATGCCGGTGTCTTCTTC.....TGATGAAGTGTCTTAATAG 483

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_29Jan04.*

1: Geneseqn1980s.*

2: Geneseqn1990s.*

3: Geneseqn2000s.*

4: Geneseqn2001as.*

5: Geneseqn2001bs.*

6: Geneseqn2002s.*

7: Geneseqn2003as.*

8: Geneseqn2003bs.*

9: Geneseqn2003cs.*

10: Geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	483	100.0	483	2	AAT12707	HCV E1 CO
2	483	100.0	483	6	AAL48916	Hepatitis
3	483	100.0	483	9	ADD5517	Hepatitis
4	474.2	98.2	480	2	AAT12708	HCV E1 CO
5	474.2	98.2	480	6	AAL48917	Hepatitis
6	474.2	98.2	480	9	ADD5519	Hepatitis
7	435	90.1	1880	2	AQ24467	NANB hepa
8	435	90.1	2187	2	ABA03491	Cuticle p
9	435	90.1	2540	2	AQ43889	NANB hepa
10	435	90.1	2540	2	AQ43753	NANBHV ge
11	433.4	89.7	1863	2	AQ063753	Fragment
12	431.8	89.4	2540	2	AQ29628	Hepatitis
13	430.2	89.1	9595	2	AAX24843	Infection
14	430.2	89.1	9595	4	AAF23492	Infection
15	430.2	89.1	9595	4	AAC86939	Nucleotid
16	430.2	89.1	9599	2	AAX24833	Infection
17	428.6	88.7	577	2	AQ35081	HCV envel
18	428.6	88.7	2187	2	ABA03492	Cuticle p
19	425.8	88.2	580	2	AQ11076	Fragment
20	425.8	88.2	580	2	AZ07647	HCV J1 B
21	425.4	88.1	577	2	AQ35085	HCV envel
22	425.4	88.1	580	2	AQ79750	Hepatitis
23	425.4	88.1	580	2	AAX00401	Hepatitis

24	425.4	88.1	580	2	AA26733	Consensus
25	425.4	88.1	580	2	AA26728	Consensus
26	425.4	88.1	1249	2	AAQ79772	Hepatitis
27	425.4	88.1	1249	2	AA26739	Consensus
28	425.4	88.1	1562	2	AAV60672	Fragment
29	425.4	88.1	1953	7	AAL55222	Plasmid p
30	425.4	88.1	2829	2	AAV60673	Fragment
31	425.4	88.1	3401	2	AAQ64069	Non-A, no
32	425.4	88.1	3401	2	AAT30387	5'UTR/COR
33	425.4	88.1	3461	2	AAQ64068	Non-A, no
34	425.4	88.1	3461	2	AAT30386	5'UTR/COR
35	425.2	88.0	567	2	AAQ27160	NANB hepa
36	424.8	88.0	642	2	AAT12704	HCV E1 CO
37	424.8	88.0	642	6	AAL48913	Hepatitis
38	424.8	88.0	642	9	ADD55634	Hepatitis
39	424.4	87.9	1270	2	AAV60668	Fragment
40	423.8	87.7	577	2	AAQ35086	HCV envel
41	423.8	87.7	580	2	AAQ79758	Hepatitis
42	423.8	87.7	580	2	AAQ79760	Hepatitis
43	423.8	87.7	580	2	AAQ79759	Hepatitis
44	423.8	87.7	795	2	AAT12705	HCV E1 CO
45	423.8	87.7	795	6	AAL48914	Hepatitis

ALIGNMENTS

RESULT 1

AAT12707
ID AAT12707 standard; DNA; 483 BP.

XX AAT12707;

DT 23-SEP-1996 (first entry)

DE HCV E1 construct HCC112A.

XX HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human;
XX serotype; reversed phase hybridisation assay; genotype; antigen; sera;
XX ss.

OS Hepatitis C virus.

PN WO9604385-A2.

XX 15-FEB-1996.

PP 31-JUL-1995; 95WO-EP003031.

PR 29-JUL-1994; 94EP-00870132.

XX (INNO-) INNOGENETICS NV.

PI Maertens G, Bosman P, De Martynoff G, Buysse M;

XX WPI; 1996-129401/13.

PT Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope proteins
- in presence of di: sulphide bond cleavage agent, to produce proteins
suitable for direct use in vaccines or diagnostic assays of HCV.

XX Claim 23; Fig 21; 146pp; English.

XX AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1
and E2 protein coding sequence constructs. These sequences are included
in vectors for the production of recombinant E1, E2, and E1/E2 proteins.
The recombinant proteins can then be isolated using a method of the
invention. In the method, the envelope proteins are purified by carrying
out a disulphide bond cleavage, or a reduction step with a disulphide
bond cleavage agent, after lysis of recombinant host cells. The
constructs containing the purified HCV envelope proteins can be used for
vaccinating humans against HCV, for in vitro detection of HCV antibodies
in a sample, and in a serotyping assay for detecting one or more

RESULT 8

ABAO3491
ID ABA03491 standard; DNA; 2187 BP.
XX
AC ABA03491;
XX
XX 15-MAR-2002 (first entry)
XX
XX Cuticle protein 1 and 2 secreting hepatitis C virus related DNA #1.
XX
XX Cuticle protein 1; cuticle protein 2; hepatitis C virus; ds.
XX
XX Unidentified.
XX
XX Key Location/Qualifiers
CDS 1..2187
FT /*tag= a
FT /product= "AAM47264"
FT /partial
FT /note= "no stop codon"
XX
XX KR97065713-A.
XX
XX 13-OCT-1997.
XX
XX 19-MAR-1996; 96KR-00007404.
XX
XX 19-MAR-1996; 96KR-00007404.
XX
XX (GLDS) LG CHEM LTD.
XX
XX Choo SH, Lee IH, Ryoo WS;
XX
XX WPI; 1998-492654/42.
XX
XX P-PSDB; AAM47264.
XX
XX Cuticle protein 1 and 2 secreting hepatitis C virus (Japanese).
XX
XX Disclosure; Page 2-4; 7pp; Korean.
XX
XX The present invention relates to cuticle protein 1 and 2 secreting
CC hepatitis C virus. The present sequence is a coding sequence provided in
CC the exemplification of the invention
XX
XX Sequence 2187 BP; 406 A; 669 C; 631 G; 481 T; 0 U; 0 Other;

Query Match 90.1%; Score 435; DB 2; Length 2187;
Best Local Similarity 94.7%; Pred. No. 7.4e-117;
Matches 450; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 2 TGCCCGGTGCTCTTCTCTATCTTCCCTTGGCCCTGCTGCTGACCATACCG 61
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 62 CTTCGGCTTATGAAGTGGCGCAACGAGTGTCCGGGTGTACCATGTGACGACTGCTCCA 121
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 566 CTTCGGCTTATGAAGTGGCGCAACGAGTGTCCGGGTGTACCATGTGACGACTGCTCCA 625
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 122 ACTCAAGCATATGTATGAGGCGAGCGACATGATGACACACCCCGGGTGGCGCT 181
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 626 ACTCAAGCATATGTATGAGGCGAGCGACATGATGACATCTCCCGGTGGCGCT 685
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 182 GCGTTCGGGAGGCGCAACTCTCCCGTGTGGGTGGGTGCTGCTGCTGCGGCGCA 241
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 686 GCGTTCGGGAGGCGCAACTCTCCCGTGTGGGTGGGTGCTGCTGCTGCGGCGCA 745
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 242 GGAAGCGCAGGCTCCCAACGACAAATACGACGACAGTGTGCTGCTGCTGCGGCTG 301
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 746 GGAATGCGAGGCTCCCACTACGACAAATACGACGACAGTGTGCTGCTGCTGCGGCTG 805
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 302 CTGCTTTCTGTTCCGCTATGATGAGTGGGGGATCTCTGCGGATCTGTTTCTGTTTCCC 361
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 806 CTGCTTTCTGTTCCGCTATGATGAGTGGGGGATCTCTGCGGATCTGTTTCTGTTTCCC 865
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 362 AGCTGTTCACTTCTCCTGCGGCGATCAACAGTACAGGACTGCACTGCTCAATCT 421
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 866 AGCTGTTCACTTCTCCTGCGGCGATCAACAGTACAGGACTGCACTGCTCAATCT 925
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 422 ATCCCGGCGCATGATACAGTCAACGCGATGCTTGGGATATGATGATGAATGCTGTC 476
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 926 ATCCCGGCGCATGATACAGTCAACGCGATGCTTGGGATATGATGATGAATGCTGTC 980
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 9

AAQ43889
ID AAQ43889 standard; cDNA to mRNA; 2540 BP.
XX
XX AAQ43889;
XX
XX 21-OCT-1993 (first entry)
XX
XX NANB hepatitis virus polynucleotide N-2540-2.
XX
XX Non-A, non-B; virus; polymerase chain reaction; detection; sensitive;
KW specific; HCV; NANBH; ss.
XX
XX Non-A.
OS non-B hepatitis virus.
XX
XX Key Location/Qualifiers
5'UTR 1..341
FT /*tag= b
FT /note= "from 5' terminal of NANBH virus RNA"
FT CDS 342..2540
FT /*tag= a
XX
XX JP05091884-A.
XX
XX 16-APR-1993.
XX
XX 10-APR-1991; 91JP-00196175.
XX
XX 12-JUN-1990; 90JP-00153401.
XX
XX 08-NOV-1990; 90JP-00304405.
XX
XX (NAKA/) NAKAMURA T.
XX
XX WPI; 1993-199637/25.
XX
XX P-PSDB; AAR38279.

Antigen related to non-A and non-B hepatitis virus - comprises non-
translation region comprising 340 - 341 mols. of nucleotides, non-
translation region comprising 1885 - 2551 mols. of nucleotides including
region 1,149 and, etc.
XX
XX Claim 3; Page 19-20; 73pp; Japanese.

The sequence is that of NANB hepatitis virus polynucleotide N-2540-2
CC which codes for a non-A, non-B (NANB) hepatitis virus gene HC-OM. The
CC polypeptide it encodes may be used in a system for detecting NANB
CC hepatitis. This method is highly specific and sensitive, and can detect
CC NANB hepatitis virus which could not be detected by conventional methods
XX

Query Match 90.1%; Score 435; DB 2; Length 2540;
Best Local Similarity 94.7%; Pred. No. 7.8e-117;
Matches 450; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 2 TGCCCGGTGCTCTTCTCTATCTTCCCTTGGCCCTGCTGCTGACCATACCG 61
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 847 TGCCCGGTGCTCTTCTCTATCTTCCCTTGGCCCTGCTGCTGACCATACCG 906
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 62 CTTCGGCTTATGAAGTGGCGCAACGAGTGTCCGGGTGTACCATGTGACGACTGCTCCA 121
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 907 CTTCGGCTTATGAAGTGGCGCAACGAGTGTCCGGGTGTACCATGTGACGACTGCTCCA 966
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||

XX HCV; infectious clone; infection; diagnosis; therapy; vaccine; screening;
 KW assay; antiviral; virucide; ss.
 XX Hepatitis C virus.
 FT Key Location/Qualifiers
 FT CDS 342..9374
 FT /*tag= a
 XX WO9904008-A2.
 XX 28-JAN-1999.
 XX 16-JUL-1998; 98WO-US014688.
 XX 18-JUL-1997; 97US-0053062P.
 XX 27-JAN-1998; 98US-00014416.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX PA
 XX PI Yanagi M, Bukh J, Emerson SU, Purcell RH;
 XX WPI; 1999-132252/11.
 XX P-PSDB; AAW98022.
 XX New isolated hepatitis C virus nucleic acids - used to develop products
 PT for the diagnosis, prevention and treatment of HCV infections and for
 PT developing screening assays.
 XX Claim 3; Fig 14A-F; 126pp; English.
 XX The present sequence comprises the nucleic acid sequence of the genome of
 CC infectious hepatitis C virus (HCV) genotype 1b strain HC-J4 (ATCC 209596)
 CC that is capable of expressing this virus when transfected into cells. HC-
 CC J4 was obtained from acute phase plasma of a chimpanzee experimentally
 CC infected with serum containing HC-J4/91. The claimed infectious nucleic
 CC acid sequence can be used to produce chimeric genomes (see AAX24833)
 CC consisting of the open reading frames of infectious nucleic acid
 CC sequences of other genotypes (including genotypes 1-6) and subtypes (such
 CC as 1b, 2a, 2b, 2c, 3a, 4a-f, 5a and 6a) of HCV. The invention also
 CC relates to the introduction of mutations or deletions into infectious
 CC nucleic acid sequences in order to produce an attenuated HCV virus
 CC suitable for vaccine development. Infectious nucleic acid sequences can
 CC also be used to produce attenuated virus via passage in vitro or in vivo
 CC of the viruses produced by transfection of a host cell with the
 CC infectious nucleic acid sequence. Vaccines comprising one or more
 CC polypeptides made from the infectious nucleic acid sequence are used to
 CC immunise mammals, especially humans, against hepatitis C. The nucleic
 CC acid sequences can also be used to induce protective immunity against the
 CC virus. The nucleic acid sequences or their encoded proteases (e.g. NS3
 CC protease) can additionally be used to develop screening assays to
 CC identify antiviral agents for HCV
 XX
 SQ Sequence 9595 BP; 1934 A; 2842 C; 2698 G; 2121 T; 0 U; 0 Other;
 Query Match 89.1%; Score 430.2; DB 2; Length 9595;
 Best Local Similarity 94.1%; Pred. No. 3.2e-115;
 Matches 447; Conservative 0; Mismatches 28; Indels 0; Gaps 0;
 QY 2 TGCCCGGTGCTCTTTCTATCTCTTCTGCGCCCTGCTGCTCTGACCATACCAG 61
 DB 847 TGCCCGGTGCTCTTTCTATCTCTTCTGCTCTGCTGCTCTGTTGACCATCCAG 906
 QY 62 CTTCCCGTTTGAAGTGGCGCAACGTGTCGGGGGTGACCATGTCAGACGACTGTCTCCA 121
 DB 907 CTTCCCGTTTGAAGTGGCGCAACGTGTCGGGGGTGACCATGTCAGACGACTGTCTCCA 966
 QY 122 ACTCAAGCATTTGTTATGAGGACGCGAGCGTATGATCATGACACCCCGGGTGGTCCCT 181
 DB 967 ACTCAAGCATTTGTTATGAGGACGCGAGCGTATGATCATGACATCTCCCGGGTGGTCCCT 1026
 QY 182 GCGTTTCGGGAGGGCAACTCTCTCCGTTGCTGGGTGGGGTCTCACTCCACGCTCGCGGCCA 241

DB 1027 GTGTTAGGAGGGTAACAGCTCCCGTTGCTGGGTAGCGCTCACTCCACGCTCGCGGCCA 1086
 QY 242 GGAAGCCAGCGTCCCCACACGACACATACGACGCCAGTCGATTTGCTGCTGGGGCTG 301
 DB 1087 GGAATGCCAGCGTCCCCACACGACACATACGACGCCAGTCGATTTGCTGCTGGGACGG 1146
 QY 302 CTGCTTTCTGTTCCGCTATGTACGTGGGGGATCTCTCGGGATCTGTTTCTTGTTCCTCC 361
 DB 1147 CTGCTTTCTGCTCGCTATGTACGTGGGGGATCTCTCGGGATCTATTTCTGCTGCTCC 1206
 QY 362 AGCTGTTACCTTCTCACTCGCGCGGATCAACAGTACAGGACTGCACTGCTCAATCT 421
 DB 1207 AGCTGTTACCTTCTCGCTCGCGCGGATGAGACAGTGCAGGACTGCAACTGCTCAATCT 1266
 QY 422 ATCCCGGCCCATGTATCAGGTCAACGCGATGCTTGGGATATGATGATGATGATGATGAT 476
 DB 1267 ATCCCGGCCCATGTATCAGGTCAACGCGATGCTTGGGATATGATGATGATGATGATGAT 1321
 RESULT 14
 AAP23492
 ID AAP23492 standard; DNA; 9595 BP.
 XX AC
 XX AAP23492;
 XX 21-MAR-2001 (first entry)
 DT Infectious Hepatitis C virus 1b genotype.
 XX DE
 XX KW GBV-B; hepatitis C virus; HCV; vaccine; ds.
 XX OS Hepatitis C virus.
 XX PN WO200075337-A1.
 XX PD 14-DEC-2000.
 XX PP 02-JUN-2000; 2000WO-US015293.
 XX PR 04-JUN-1999; 99US-0137694P.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX PI Bukh J, Yanagi M, Emerson SU, Purcell RH;
 XX WPI; 2001-091214/10.
 XX New infectious nucleic acids of the GB virus-B clone, useful for
 PT indirectly studying the molecular properties of hepatitis C virus (HCV)
 PT and in developing vaccines and therapeutics for HCV.
 XX PS Disclosure; Fig 7; 96pp; English.
 XX
 SQ Sequence 9595 BP; 1934 A; 2843 C; 2697 G; 2121 T; 0 U; 0 Other;
 Query Match 89.1%; Score 430.2; DB 4; Length 9595;
 Best Local Similarity 94.1%; Pred. No. 3.2e-115;
 Matches 447; Conservative 0; Mismatches 28; Indels 0; Gaps 0;
 QY 2 TGCCCGGTGCTCTTTCTATCTCTTCTGCGCCCTGCTGCTCTGACCATACCAG 61
 DB 847 TGCCCGGTGCTCTTTCTATCTCTTCTGCTCTGCTGCTCTGTTGACCATCCAG 906
 QY 62 CTTCCCGTTTGAAGTGGCGCAACGTGTCGGGGGTGACCATGTCAGACGACTGTCTCCA 121
 DB 907 CTTCCCGTTTGAAGTGGCGCAACGTGTCGGGGGTGACCATGTCAGACGACTGTCTCCA 966

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2004, 09:12:53 ; Search time 1298.46 Seconds
(without alignments)
11108.065 Million cell updates/sec

Title: US-09-899-303a-9
Perfect score: 483
Sequence: 1 ATGCCCGGTGCTCTTCTC.....TGATGAACGTGCTCTTAATAG 483

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*

1: em_estba.*
2: em_esthum.*

3: em_estin.*
4: em_estmu.*

5: em_estov.*
6: em_estpl.*

7: em_estro.*
8: em_hic.*

9: gb_est1.*
10: gb_est2.*

11: gb_hic.*
12: gb_est3.*

13: gb_est4.*
14: gb_est5.*

15: em_estfun.*
16: em_estom.*

17: em_gss_hum.*
18: em_gss_in.*

19: em_gss_pin.*
20: em_gss_vrt.*

21: em_gss_fun.*
22: em_gss_mam.*

23: em_gss_mus.*
24: em_gss_pro.*

25: em_gss_rod.*
26: em_gss_pbg.*

27: em_gss_vrl.*
28: gb_gss1.*

29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	67.2	13.9	488	9	AV755731
C 2	55.6	11.5	492	9	AV758366
C 3	43.2	8.9	534	14	CF846043
C 4	41.6	8.6	664	28	BZ645446

C 5	41.6	8.6	741	29	CG300832
C 6	41.6	8.6	794	29	CG213867
C 7	41.6	8.6	995	29	CC634840
C 8	40.8	8.4	761	29	CG373319
C 9	40.5	8.4	526	9	AL825643
C 10	40.2	8.3	399	9	AV638521
C 11	40.2	8.3	434	9	AV637507
C 12	40.2	8.3	440	9	AV637983
C 13	40.2	8.3	450	9	AV637259
C 14	40.2	8.3	451	9	AV637328
C 15	40.2	8.3	451	9	AV637643
C 16	40.2	8.3	453	9	AV634724
C 17	40.2	8.3	454	9	AV637050
C 18	40.2	8.3	456	9	AV635382
C 19	40.2	8.3	473	9	AV632765
C 20	40.2	8.3	481	9	AV635503
C 21	40.2	8.3	485	9	AV632811
C 22	40.2	8.3	506	9	AV392445
C 23	40.2	8.3	508	9	AV634095
C 24	40.2	8.3	526	9	AV641895
C 25	40.2	8.3	533	9	AV638125
C 26	40.2	8.3	537	9	AV632335
C 27	40.2	8.3	588	9	AV387329
C 28	40.2	8.3	983	29	CG634833
C 29	40.2	8.3	1186	13	EX421743
C 30	39.4	8.2	1201	9	AL565958
C 31	39.2	8.1	624	14	CD206870
C 32	39.2	8.1	656	14	CB924688
C 33	39.2	8.1	497	9	AV633658
C 34	39.2	8.1	610	14	CB657655
C 35	39.2	8.1	856	28	BZ578381
C 36	39.2	8.1	872	28	BZ555011
C 37	38.4	8.0	645	29	CNS01213
C 38	38.4	8.0	734	29	CC916826
C 39	38.4	8.0	742	29	CC734434
C 40	38.4	8.0	771	28	BZ530934
C 41	38.4	8.0	838	29	CC734426
C 42	37.8	7.8	490	9	AV634529
C 43	37.8	7.8	705	14	CA618797
C 44	37.8	7.8	1039	13	EX415186
C 45	37.6	7.8	309	12	BI098866

ALIGNMENTS

RESULT 1

AV755731/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

AV755731 488 bp mRNA linear EST 19-OCT-2000
AV755731 BM Homo sapiens cDNA clone BMFAKB03 5', mRNA sequence.
AV755731
AV755731.1 GI:10913579

EST.

Homo sapiens (human)

Homo sapiens

Bukarya; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 488)

Gu, J., Zhao, M., Huang, Q., Xu, X., Li, Y., Peng, Y., Song, H., Xiao, H.,

Gu, Y., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Chang, Z., Xu, Z.,

Zeng, L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M.,

Lu, G., Yang, Y., Gao, G., Wang, Z., Zhang, Q., Chen, S., Han, Z. and

Chen, Z.

Homo sapiens cDNA BM clones

Unpublished (2000)

Contact: Zeguang Han

Chinese National Human Genome Center at Shanghai

351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai

201203, P. R. China

Tel: 86-21-50801919(ex.45)

Fax: 86-21-50801922

Email: hanzg@chgc.sh.cn

This clone is available at CHGC in Shanghai.

```

FEATURES
  source
    Location/Qualifiers
      1..488
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="BMFAK303"
        /tissue_type="Bone marrow"
        /cell_type="CD34+ hematopoietic stem/progenitor cell"
        /lab_host="BM25.8"
        /clone_lib="BM"
        /note="Vector: pTriplex2; Site_1: sfIIA; Site_2: sfIIB"

ORIGIN
  Query Match      13.9%; Score 67.2; DB 9; Length 488;
  Best Local Similarity 67.0%; Pred. No. 2.8e-06;
  Matches 126; Conservative 0; Mismatches 58; Indels 4; Gaps 2;

QY 292 GTTGGGCTGCTGTTCTGTTCCGCTATGTAAGTGGGGATCTCTGCGGATCTGTTTC 351
DB 472 GTGGGTGCACACTCGCTCTCTGCTCAGCTCTCTGAGTGGGACCTCTGCGACGGAGTGATG 413
QY 352 CTGTTTCCAGCTGTTACCTTCTCACCTCCGCGCATCAACAGTACAGGACTGCAAC 411
DB 412 CTTCAGTTCAGCTATCA---TCTGGCTCAGACCAATGAGTTGTGATGATGCAAC 356
QY 412 TGCTCAATCTATCCCGCCCATGATCAGGTACCGCATG-GCTTGGGATATGATGATGAA 470
DB 355 TGCTCCATCTATCTCTGGCCCATCACTGGACACCGTATGAGCATGGGACATGATGATGAA 296
QY 471 CTGGTCTCT 478
DB 295 CTGGTCTG 288

RESULT 2
AV758366/C
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
  Homo sapiens (human)
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 492)
  Gu,J., Zhao,M., Huang,Q., Xu,X., Li,Y., Peng,Y., Song,H., Xiao,H.,
  Zeng,L., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M.,
  Lu,G., Yang,Y., Gao,G., Wang,Z., Zhang,Q., Chen,S., Han,Z. and
  Chen,Z.
  Homo sapiens cDNA BM clones
  Unpublished (2000)
  Contact: Zeguang Han
  Chinese National Human Genome Center at Shanghai
  351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
  201203, P. R. China
  Tel: 86-21-50801919 (ex.45)
  Fax: 86-21-50801922
  Email: hanzg@chgc.sh.cn
  This clone is available at CHGC in Shanghai.
  Location/Qualifiers
    1..492
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="BMFAK303"
      /tissue_type="Bone marrow"
      /cell_type="CD34+ hematopoietic stem/progenitor cell"
      /lab_host="BM25.8"
      /clone_lib="BM"
      /note="Vector: pTriplex2; Site_1: sfIIA; Site_2: sfIIB"

ORIGIN
  Query Match      8.9%; Score 43.2; DB 14; Length 534;
  Best Local Similarity 47.4%; Pred. No. 3.4;
  Matches 129; Conservative 0; Mismatches 143; Indels 0; Gaps 0;

QY 70 TATGAAGTGGCAACGTTCCGGGGTGTACCATGTACAGNACGACTGCTCAACTCAGC 129
DB 200 TAGGGCGTGGCGGAGATTTCAGGTATCCGATGGGCTTCGCGGCTTCTACAACTGGAC 259
QY 130 ATAGTGTATGAGGACGACGACATGATGCACACCCCCCGGGTGGCGCTTGGTTGG 189
DB 260 TCGATGGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 319

```

```

FEATURES
  source
    Location/Qualifiers
      1..534
        /organism="Phytophthora sojae"
        /mol_type="mRNA"
        /db_xref="taxon:67593"
        /clone="sHB036B09"
        /tissue_type="mycelium"
        /cell_line="P6497"
        /dev_stage="48 hr. post infection stage"
        /lab_host="Soybean plant"
        /clone_lib="USDA-IPAFS:Expression of Phytophthora sojae
        genes during infection and propagation_sHB"
        /note="Vector: pBK-CMV, Site_1: EcoRI; Site_2: XhoI"

ORIGIN
  Query Match      8.9%; Score 43.2; DB 14; Length 534;
  Best Local Similarity 47.4%; Pred. No. 3.4;
  Matches 129; Conservative 0; Mismatches 143; Indels 0; Gaps 0;

QY 70 TATGAAGTGGCAACGTTCCGGGGTGTACCATGTACAGNACGACTGCTCAACTCAGC 129
DB 200 TAGGGCGTGGCGGAGATTTCAGGTATCCGATGGGCTTCGCGGCTTCTACAACTGGAC 259
QY 130 ATAGTGTATGAGGACGACGACATGATGCACACCCCCCGGGTGGCGCTTGGTTGG 189
DB 260 TCGATGGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 319

Query Match      11.5%; Score 55.6; DB 9; Length 492;
Best Local Similarity 65.9%; Pred. No. 0.0024;
Matches 112; Conservative 0; Mismatches 54; Indels 4; Gaps 2;

QY 308 TCTGTTCCGCTATGTACGTTGGGGGATCTCTGCGGATCTGTTTCTTCCAGCTCT 367
DB 457 TGTGATCAGCTACTACGTTGTGGACCTCTGCGTTGGGATATCGCTTGAGCCCACTGA 398
QY 368 TCACCTTCTTCACTCGCCGGCATCAACAGTACAGGACTGCAACTGCTCAATCTATCCCG 427
DB 397 TTA---TCTCTCAGCAGCAACATGTTGTGCAAGAATGCAACTGCTATTCTATCTG 341
QY 428 GCATGTATCAGGTC-ACCGATGCTTGGGATATGATGATGATGATGATGATGATGATG 476
DB 340 GCTGCATCATCGCTACAGTATGCGATAGGCTATGATGATGATGATGATGATGATGATG 291

RESULT 3
CF846043
LOCUS
DEFINITION
  534 bp mRNA linear EST 30-OCT-2003
  pHB036B09f USDA-IPAFS:Expression of phytophthora sojae genes
  during infection and propagation_sHB Phytophthora sojae cDNA clone
  sHB036B09 5, mRNA sequence.
ACCESSION
  CF846043
  GI:38061697
VERSION
  EST.
KEYWORDS
  Phytophthora sojae
SOURCE
  Phytophthora sojae
  Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae;
  Phytophthora.
REFERENCE
  1 (bases 1 to 534)
  Tyler,B.
  Tyler,B. Not Published
  Unpublished (2003)
  Contact: Tyler B
  Tyler lab
  VBI
  1880 Pratt Dr., Blacksburg, VA 24061, USA
  Tel: 540-231-7318
  Email: bmtyle@vt.edu
  PCR Primers
  FORWARD: BK reverse primer
  BACKWARD: BK reverse primer
  Plate: 036 row: B column: 09
  Seq primer: BK reverse primer
  High quality sequence stop: 534.
  Location/Qualifiers
    1..534
      /organism="Phytophthora sojae"
      /mol_type="mRNA"
      /db_xref="taxon:67593"
      /clone="sHB036B09"
      /tissue_type="mycelium"
      /cell_line="P6497"
      /dev_stage="48 hr. post infection stage"
      /lab_host="Soybean plant"
      /clone_lib="USDA-IPAFS:Expression of Phytophthora sojae
      genes during infection and propagation_sHB"
      /note="Vector: pBK-CMV, Site_1: EcoRI; Site_2: XhoI"

ORIGIN
  Query Match      8.9%; Score 43.2; DB 14; Length 534;
  Best Local Similarity 47.4%; Pred. No. 3.4;
  Matches 129; Conservative 0; Mismatches 143; Indels 0; Gaps 0;

QY 70 TATGAAGTGGCAACGTTCCGGGGTGTACCATGTACAGNACGACTGCTCAACTCAGC 129
DB 200 TAGGGCGTGGCGGAGATTTCAGGTATCCGATGGGCTTCGCGGCTTCTACAACTGGAC 259
QY 130 ATAGTGTATGAGGACGACGACATGATGCACACCCCCCGGGTGGCGCTTGGTTGG 189
DB 260 TCGATGGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 319

```

Qy	190	GAGGGAACTCTCCGTTGCTGGTGGCGCTCACTCCACGCTCGCGCGCAGGACGCC	249
Db	320	CRTACGGCGGCACCATCTCGGCTCGAACGCTGGTCTCGACGTGCGACGANGATTATC	379
Qy	250	AGCTCCGCCACAAAGCAATACGACGCCACGTCGATTTGCTGTTGGGGCTGCTGCTTTC	309
Db	380	AAC TTC TCGACGACGAACGGCGCTCTCGCAGGTATACGTGATCGCGCGTGAACGGCACCCAC	439
Qy	310	TGTTCCGCTATGTACGTGGGGGATCTCTGCGG	341
Db	440	CGTCCGCGCCAAAGATCTCGGAGGAGTCCGCG	471

RESULT 4	BZ645446	664 bp	linear	GSS 29-JAN-2003
LOCUS	BZ645446/c			
DEFINITION	OCB386TC ZM_0.7_1.5_KB Zea mays genomic clone ZMWBMa013P04,			
	genomic survey sequence.			
ACCESSION	BZ645446			
VERSION	BZ645446.1			
KEYWORDS	GI:28107610			
	GSS.			

```

FEATURES
source
Class: smeared ends.
Location/Qualifiers
1..664
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZM4BMA013P04"
/clone_lib="ZM_0.7_1.5_KB"
/note="vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

```

RESULT 5	CG300832/c
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	

ORIGIN	Query Match Best Loc Matches
QY	
Db	
QY	
Db	
QY	
Db	
QY	
Db	

RESULT 6	ACCESSION
CG213867/c	VERSION
LOCUS	KEYWORDS
DEFINITION	SOURCE
	ORGANISM

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 794)
Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.

Consortium for Maize Genomics

Unpublished (2002)

Other GSSs: OGVCR53TH

Contact: Cathy Whitelaw

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843

Fax: 301-838-0208

Email: whitelaw@tigr.org

Seq primer: TF

Class: sheared ends.

Location/Qualifiers

1..794

/organism="Zea mays"

/mol_type="genomic DNA"

/strain="B73"

/db_xref="taxon:4577"

/clone="ZMMBma0506110"

/clone_lib="ZM_0.7_1.5_KB"

/note="Vector: pBCSK; Site 1: HincII; 0.7-1.5 kb

methylation filtered genomic DNA library"

ORIGIN

Query Match 8.6%; Score 41.6; DB 29; Length 794;
Best Local Similarity 50.5%; Pred. No. 10;
Matches 101; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

QY 119 CCAACTCAAGCATAGTGTATGAGGCGAGCATATCATGCACACCCCGGTGGTGC 178
DB 397 CGACGACGACACGTCGTGGTGGCCCAAGAGCGAGTGGGAGCGGGGCCGAC 338
QY 179 CTTGGTTCGGGAGGCAACTCTCCCGTGTGGTGGTGGTGGTGGTGGTGGTGG 238
DB 337 CGTACTCGGTGGGAGCGCTCCAGAGTGTAGAGCCCTGCGCGTGTGGGG 278
QY 239 CCAGAACGCGAGCTCCCAAGACACATAGAGCGGCGATGCTGATTTGGTGGG 298
DB 277 ACAGGAGCGGCGGCTCTCCAGCGGACCAACAGCGGCTACCCCTTCTCGGTGG 218
QY 299 CTGCTGCTTCTGTTCCGCT 318
DB 217 GTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 198

RESULT 7

CC634840/c
LOCUS
DEFINITION
OGVCR40TV ZM_0.7_1.5_KB Zea mays genomic clone ZMMBma050608,
genomic survey sequence.

ACCESSION
CC634840

VERSION
CC634840.1 GI:32012237

KEYWORDS
GSS.

SOURCE
Zea mays

ORGANISM
Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 995)
Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.

Consortium for Maize Genomics

Unpublished (2002)

Other GSSs: OGVCR40TH

Contact: Cathy Whitelaw

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843

Fax: 301-838-0208

Email: whitelaw@tigr.org

Seq primer: TF

Class: sheared ends.

Location/Qualifiers

1..995

/organism="Zea mays"

/mol_type="genomic DNA"

/strain="B73"

/db_xref="taxon:4577"

/clone="ZMMBma050608"

/clone_lib="ZM_0.7_1.5_KB"

/note="Vector: pBCSK; Site 1: HincII; 0.7-1.5 kb

methylation filtered genomic DNA library"

ORIGIN

Query Match 8.6%; Score 41.6; DB 29; Length 995;
Best Local Similarity 50.5%; Pred. No. 11;
Matches 101; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

QY 119 CCAACTCAAGCATAGTGTATGAGGCGAGCATATCATGCACACCCCGGTGGTGC 178
DB 864 CGACGACGACACGTCGTGGTGGCCCAAGAGCGAGTGGGAGCGGGGCCGAC 805
QY 179 CTTGGTTCGGGAGGCAACTCTCCCGTGTGGTGGTGGTGGTGGTGGTGGTGG 238
DB 804 CGTACTCGGTGGGAGCGCTCCAGAGTGTAGAGCCCTGCGCGTGTGGTGGCG 745
QY 239 CCAGAACGCGAGCTCCCAAGACACATAGAGCGGCGATGCTGATTTGGTGGG 298
DB 744 ACAGGAGCGGCGGCTCTCCAGCGGACCAACAGCGGCTACCCCTTCTCGGTGG 685
QY 299 CTGCTGCTTCTGTTCCGCT 318
DB 684 GTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 565

RESULT 8

CG373319/c

LOCUS

DEFINITION

CG373319

ACCESSION

VERSION

CG373319.1 GI:34290586

KEYWORDS

SOURCE

ORGANISM

Zea mays

Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

clade; Panicoideae; Andropogoneae; Zea.

REFERENCE

AUTHORS

Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,

Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,

Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.

Consortium for Maize Genomics

Unpublished (2002)

Other GSSs: OG1BI44TH

Contact: Cathy Whitelaw

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843

Fax: 301-838-0208

Email: whitelaw@tigr.org

Seq primer: TF

Class: sheared ends.

Location/Qualifiers

1..761

/organism="Zea mays"

/mol_type="genomic DNA"

/strain="B73"

/db_xref="taxon:4577"

/clone="ZMMBma0723H15"


```

ACCESSION      AV637507.1      GI:10780827
VERSION        AV637507.1
KEYWORDS       EST.
SOURCE         Chlamydomonas reinhardtii
ORGANISM       Chlamydomonas reinhardtii
REFERENCE      Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
AUTHORS       Chlamydomonas reinhardtii
              1 (bases 1 to 434)
              Asamizu, E., Miura, K., Kucho, K., Inoue, Y., Fukuzawa, H., Ohyama, K.,
              Nakamura, Y. and Tabata, S.
TITLE          Generation of expressed sequence tags from low-CO2 and high-CO2
              adapted cells of Chlamydomonas reinhardtii
JOURNAL        DNA Res. 7 (5), 305-307 (2000)
MEDLINE        20539644
PUBMED        11089912
COMMENT        Contact: Erika Asamizu
              The First Laboratory for Plant Gene Research
              Kazusa DNA Research Institute
              Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
              Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
FEATURES       Location/Qualifiers
              1..434
              /organism="Chlamydomonas reinhardtii"
              /mol_type="mRNA"
              /strain="C9"
              /db_xref="taxon:3055"
              /clone_lib="HC074a01_r"
              /note="Vector: pBluescriptII SK-; Site 1: EcoRI; Site 2:
              XhoI; The cDNA library was constructed from cells cultured
              in a medium with bubbling air containing 5% carbon
              dioxide"
ORIGIN
Query Match      8.3%; Score 40.2; DB 9; Length 434;
Best Local Similarity 51.4%; Pred. No. 18;
Matches 93; Conservative 0; Mismatches 88; Indels 0; Gaps 0;
QY      30 CTTGGCCCTGCTGCTCTGACCATACAGCTTCGCTTATGAGTGCGCAAGCTGTC 89
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      21 CTTGACCCCTGGAGCGCTGCTCATCTGTCGACCACTTCGCGGTGGCACCATCTT 80
QY      90 CGGGGTGACCATGTGACGACGAGCTGCTCAACTCAAGCATAGTGTATGAGCGACCGGA 149
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      81 CGGGGTGTCGCGCGGACAGCTGACCAACATCATCCGCGCGGAGGTGCTGCCGG 140
QY      150 CATGATCATGACACACCCCGGTGCTGCTGCTTCAAGTGTGAGGAGCGACGGA 209
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      141 CATGGGCATCTACGGTCCCGCACCGCTGTTCTGCAATGCTGCAAGGACGCGCCGCGCTG 200
QY      210 C 210
DB      201 C 201

RESULT 12
LOCUS      AV637983
DEFINITION AV637983 Chlamydomonas reinhardtii 5% CO2 Chlamydomonas reinhardtii
            cDNA clone HC080c04_r 5', mRNA sequence.
ACCESSION AV637983
VERSION   AV637983
KEYWORDS  EST.
SOURCE    Chlamydomonas reinhardtii
ORGANISM  Chlamydomonas reinhardtii
REFERENCE 1 (bases 1 to 440)
AUTHORS   Asamizu, E., Miura, K., Kucho, K., Inoue, Y., Fukuzawa, H., Ohyama, K.,
            Nakamura, Y. and Tabata, S.
TITLE     Generation of expressed sequence tags from low-CO2 and high-CO2
            adapted cells of Chlamydomonas reinhardtii
JOURNAL   DNA Res. 7 (5), 305-307 (2000)

ACCESSION      AV637507.1      GI:10780827
VERSION        AV637507.1
KEYWORDS       EST.
SOURCE         Chlamydomonas reinhardtii
ORGANISM       Chlamydomonas reinhardtii
REFERENCE      Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
AUTHORS       Chlamydomonas reinhardtii
              1 (bases 1 to 434)
              Asamizu, E., Miura, K., Kucho, K., Inoue, Y., Fukuzawa, H., Ohyama, K.,
              Nakamura, Y. and Tabata, S.
TITLE          Generation of expressed sequence tags from low-CO2 and high-CO2
              adapted cells of Chlamydomonas reinhardtii
JOURNAL        DNA Res. 7 (5), 305-307 (2000)
MEDLINE        20539644
PUBMED        11089912
COMMENT        Contact: Erika Asamizu
              The First Laboratory for Plant Gene Research
              Kazusa DNA Research Institute
              Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
              Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
FEATURES       Location/Qualifiers
              1..434
              /organism="Chlamydomonas reinhardtii"
              /mol_type="mRNA"
              /strain="C9"
              /db_xref="taxon:3055"
              /clone_lib="HC074a01_r"
              /note="Vector: pBluescriptII SK-; Site 1: EcoRI; Site 2:
              XhoI; The cDNA library was constructed from cells cultured
              in a medium with bubbling air containing 5% carbon
              dioxide"
ORIGIN
Query Match      8.3%; Score 40.2; DB 9; Length 434;
Best Local Similarity 51.4%; Pred. No. 18;
Matches 93; Conservative 0; Mismatches 88; Indels 0; Gaps 0;
QY      30 CTTGGCCCTGCTGCTCTGACCATACAGCTTCGCTTATGAGTGCGCAAGCTGTC 89
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      21 CTTGACCCCTGGAGCGCTGCTCATCTGTCGACCACTTCGCGGTGGCACCATCTT 80
QY      90 CGGGGTGACCATGTGACGACGAGCTGCTCAACTCAAGCATAGTGTATGAGCGACCGGA 149
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      81 CGGGGTGTCGCGCGGACAGCTGACCAACATCATCCGCGCGGAGGTGCTGCCGG 140
QY      150 CATGATCATGACACACCCCGGTGCTGCTGCTTCAAGTGTGAGGAGCGACGGA 209
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      141 CATGGGCATCTACGGTCCCGCACCGCTGTTCTGCAATGCTGCAAGGACGCGCCGCGCTG 200
QY      210 C 210
DB      201 C 201

RESULT 13
LOCUS      AV637259
DEFINITION AV637259 Chlamydomonas reinhardtii 5% CO2 Chlamydomonas reinhardtii
            cDNA clone HC070g06_r 5', mRNA sequence.
ACCESSION AV637259
VERSION   AV637259
KEYWORDS  EST.
SOURCE    Chlamydomonas reinhardtii
ORGANISM  Chlamydomonas reinhardtii
REFERENCE 1 (bases 1 to 450)
AUTHORS   Asamizu, E., Miura, K., Kucho, K., Inoue, Y., Fukuzawa, H., Ohyama, K.,
            Nakamura, Y. and Tabata, S.
TITLE     Generation of expressed sequence tags from low-CO2 and high-CO2
            adapted cells of Chlamydomonas reinhardtii
JOURNAL   DNA Res. 7 (5), 305-307 (2000)
MEDLINE   20539644
PUBMED    11089912
COMMENT    Contact: Erika Asamizu
            The First Laboratory for Plant Gene Research
            Kazusa DNA Research Institute
            Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
            Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
FEATURES     Location/Qualifiers
              1..450
              /organism="Chlamydomonas reinhardtii"
              /mol_type="mRNA"
              /strain="C9"
              /db_xref="taxon:3055"

```

```

MEDLINE        20539644
PUBMED        11089912
COMMENT        Contact: Erika Asamizu
              The First Laboratory for Plant Gene Research
              Kazusa DNA Research Institute
              Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
              Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
FEATURES       Location/Qualifiers
              1..440
              /organism="Chlamydomonas reinhardtii"
              /mol_type="mRNA"
              /strain="C9"
              /db_xref="taxon:3055"
              /clone_lib="HC080c04_r"
              /note="Vector: pBluescriptII SK-; Site 1: EcoRI; Site 2:
              XhoI; The cDNA library was constructed from cells cultured
              in a medium with bubbling air containing 5% carbon
              dioxide"
ORIGIN
Query Match      8.3%; Score 40.2; DB 9; Length 440;
Best Local Similarity 51.4%; Pred. No. 18;
Matches 93; Conservative 0; Mismatches 88; Indels 0; Gaps 0;
QY      30 CTTGGCCCTGCTGCTCTGACCATACAGCTTCGCTTATGAGTGCGCAAGCTGTC 89
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      80 CTTGACCCCTGGAGCGCTGCTCATCTGTCGACCACTTCGCGGTGGCACCATCTT 139
QY      90 CGGGGTGACCATGTGACGACGAGCTGCTCAACTCAAGCATAGTGTATGAGCGACCGGA 149
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      140 CGGGGTGTCGCGCGGACAGCTGACCAACATCATCCGCGCGGAGGTGCTGCCGG 199
QY      150 CATGATCATGACACACCCCGGTGCTGCTGCTTCAAGTGTGAGGAGCGCAACTCTCTCCGCTG 209
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      200 CATGGGCATCTACGGTCCCGCACCGCTGTTCTGCAATGCTGCAAGGACGCGCCGCGCTG 259
QY      210 C 210
DB      260 C 260

RESULT 13
LOCUS      AV637259
DEFINITION AV637259 Chlamydomonas reinhardtii 5% CO2 Chlamydomonas reinhardtii
            cDNA clone HC070g06_r 5', mRNA sequence.
ACCESSION AV637259
VERSION   AV637259
KEYWORDS  EST.
SOURCE    Chlamydomonas reinhardtii
ORGANISM  Chlamydomonas reinhardtii
REFERENCE 1 (bases 1 to 450)
AUTHORS   Asamizu, E., Miura, K., Kucho, K., Inoue, Y., Fukuzawa, H., Ohyama, K.,
            Nakamura, Y. and Tabata, S.
TITLE     Generation of expressed sequence tags from low-CO2 and high-CO2
            adapted cells of Chlamydomonas reinhardtii
JOURNAL   DNA Res. 7 (5), 305-307 (2000)
MEDLINE   20539644
PUBMED    11089912
COMMENT    Contact: Erika Asamizu
            The First Laboratory for Plant Gene Research
            Kazusa DNA Research Institute
            Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
            Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
FEATURES     Location/Qualifiers
              1..450
              /organism="Chlamydomonas reinhardtii"
              /mol_type="mRNA"
              /strain="C9"
              /db_xref="taxon:3055"

```


Db 376 C 376

Search completed: June 16, 2004, 13:27:13
Job time : 1300.46 secs

[illegible]

RESULT 5

```

US-08-470-426B-17
; Sequence 17, Application US/08470426B
; Patent No. 5856458
; GENERAL INFORMATION:
; APPLICANT: Okamoto, Hiroaki
; APPLICANT: Nakamura, Tetsuo
; TITLE OF INVENTION: OLIGONUCLEOTIDE PRIMERS, AND THEIR
; TITLE OF INVENTION: APPLICATION FOR HIGH-FIDELITY DETECTION OF NON-A, NON-B
; TITLE OF INVENTION: HEPATITIS VIRUS
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Beveridge, DeGrandi, Weillacher & Young,
; ADDRESSBB: L.I.P.
; STREET: 1850 M Street, N.W., Suite 800
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/470,426B
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-153402
; FILING DATE: 12-JUN-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Weillacher, Robert G.
; REGISTRATION NUMBER: 20,531
; REFERENCE/DOCKET NUMBER: 06/59-47083.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2811
; TELEFAX: (202) 659-1462
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1539 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)

```



```

; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/612,973
; FILING DATE: 11-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 642 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..639
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..636
US-08-927-597-3

```

```

Query Match      88.0%; Score 424.8; DB 3; Length 642;
Best Local Similarity 93.3%; Pred. No. 2.6e-111;
Matches 444; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

Qy      1  ATGCCCGGTGCTCTTTCTATCTCTCTTGGCCCTCTCTCTGTCTGACCATACCA 60
Db      1  ATGCCCGGTGCTCTTTCTATCTCTCTTGGCTTTACTGTCTGTCTGACCATTCOA 60

Qy      61  GCTTCCGCTTATGAAGTCGCAACGTGTCGGGGTGACCATGTTCACGAACGACTGCTCC 120
Db      61  GCTTCCGCTTATGAGTGGCAACGTGTCGGGATGTACCATGTTCACGAACGACTGCTCC 120

Qy     121  AACTCAGCATGTGTATGAGCGAGCGACATGATGTCACACCCCGGGTGCGTGCC 180
Db     121  AACTCAAGCATGTGTATGAGCGAGCGACATGATGTCACACCCCGGGTGCGTGCC 180

Qy     181  TCGGTTCCGGAGGGCAACTCTCCGCTTGTGGGTGGCGCTCACCTCCACGCTCGCGGC 240
Db     181  TCGGTTCCGGAGAACACTCTTCCGCTGTGGGTAGCGCTCACCCCAAGCTCGAGCT 240

Qy     241  AGGAACGCCAGCGTCCCAACCAACGAATAACGACGCCACGTCGATTTGCTGGGGCT 300
Db     241  AGGAACGCCAGCGTCCCAACCAACGAATAACGACGCCACGTCGATTTGCTGGGGGG 300

Qy     301  GCTGCTTCTGTTCCGCTATGTACGTGGGGGATCTCTGGGATCTGTTTCTGTTTCC 360
Db     301  GCTGCTCTCTGTTCCGCTATGTACGTGGGGGATCTCTGGGATCTGTTTCTGTTTCC 360

Qy     361  CAGCTGTTCACTTCTCACTCGCGGATCAACAGTACAGGATCGCACTGCTCAATC 420
Db     361  CAGCTGTTCACTTCTCGCTCGCGGATGAGAGCGGTGCGAGGACTGCAATTCCTCAATC 420

Qy     421  TATCCCGGCCATGATCAAGGTACCGCATGGCTTGGGATGATGATGAAGTGGTC 476
Db     421  TATCCCGGCCATGATCAAGGTACCGTATGGCTTGGGATGATGATGAAGTGGTC 476

```

Search completed: June 16, 2004, 13:30:23
Job time : 35.8162 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2004, 08:32:08 ; Search time 1464.4 Seconds
(without alignments)
14206.949 Million cell updates/sec

Title: US-09-899-303A-11
Perfect score: 480
Sequence: 1 ATGTCGGTTCCTCTTCTC.....TGATGATGACTGGTAATAG 480

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum Match 100%
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb_ba.*
- 2: gb_hg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_ats.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_ats.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rtd.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sv.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	480	100.0	480	6	A48673	A48673 Sequence 11
2	480	100.0	480	6	AR157328	AR157328 Sequence
3	480	100.0	480	6	AX452760	AX452760 Sequence
4	480	100.0	480	6	AX685012	AX685012 Sequence
5	480	100.0	480	6	AX794854	AX794854 Sequence
6	474.2	98.8	483	6	A48671	A48671 Sequence 9
7	474.2	98.8	483	6	AR157327	AR157327 Sequence
8	474.2	98.8	483	6	AX452758	AX452758 Sequence
9	474.2	98.8	483	6	AX685010	AX685010 Sequence
10	474.2	98.8	483	6	AX794852	AX794852 Sequence
11	436	90.8	9379	14	AP207766	AP207766 Hepatitis
12	432.8	90.2	1880	14	HPCSTRJ4	D00832 Hepatitis C
13	432.8	90.2	2540	6	E04260	E04260 cDNA encodi
14	432.8	90.2	2540	6	E04805	E04805 cDNA to 5'
15	432.8	90.2	2540	6	E07391	E07391 cDNA encodi
16	432.8	90.2	9448	14	HPCJ483	D13558 Hepatitis C
17	431.2	89.8	1539	6	AR027786	AR027786 Sequence
18	431.2	89.8	1863	6	AR027783	AR027783 Sequence
19	430.8	89.8	9580	14	AF054250	AF054250 Hepatitis
20	429.6	89.5	8780	14	AF054257	AF054257 Hepatitis
21	429.6	89.5	8780	14	AF054258	AF054258 Hepatitis
22	428	89.2	8780	14	AF054255	AF054255 Hepatitis
23	428	89.2	8781	14	AF054254	AF054254 Hepatitis
24	428	89.2	9379	14	AF165052	AF165052 Hepatitis
25	428	89.2	9460	14	HPCJ491	D10750 Hepatitis C
26	428	89.2	9595	6	AR119832	AR119832 Sequence
27	428	89.2	9595	14	AF054247	AF054247 Hepatitis
28	428	89.2	9596	14	AF054249	AF054249 Hepatitis
29	428	89.2	9599	6	AR119833	AR119833 Sequence
30	426.4	88.8	577	6	E04085	E04085 gDNA encodi
31	426.4	88.8	618	14	HPCB1P12	D28929 Hepatitis C
32	426.4	88.8	8779	14	AF054251	AF054251 Hepatitis
33	426.4	88.8	8780	14	AF054252	AF054252 Hepatitis
34	426.4	88.8	9410	14	HPCJLR2	D50481 Hepatitis C
35	426.4	88.8	9595	14	AF054248	AF054248 Hepatitis
36	426.2	88.8	633	6	A48669	A48669 Sequence 7
37	426.2	88.8	633	6	AR157326	AR157326 Sequence
38	426.2	88.8	633	6	AX452756	AX452756 Sequence
39	426.2	88.8	633	6	AX685008	AX685008 Sequence
40	426.2	88.8	633	6	AX794850	AX794850 Sequence
41	426	88.8	9377	14	AP207756	AP207756 Hepatitis
42	426	88.8	9435	14	AB049093	AB049093 Hepatitis
43	424.8	88.5	3296	14	AB008446	AB008446 Hepatitis
44	424.8	88.5	8780	14	AF054259	AF054259 Hepatitis
45	424.8	88.5	8781	14	AF054256	AF054256 Hepatitis

ALIGNMENTS

RESULT 1
A48673
LOCUS A48673 Sequence 11 from Patent WO9604385. 480 bp DNA linear PAT 07-MAR-1997
DEFINITION A48673
ACCESSION A48673
VERSION A48673.1 GI:2302386
KEYWORDS .
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 480)
AUTHORS Maertens,G., Bosman,F., De M.G. and Buysse,M.
TITLE PURIFIED HEPATITIS C VIRUS ENVELOPE PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
JOURNAL Patent: WO 9604385-A 11 15-FEB-1996;

AYCQTRGLGCIITSLGRDNKNOVEGVVVSTATQSTFLATCVANGVCVVVYHGAGSK
 TLADGPKGPIOTWYNDVDDQGLWGLAPPGARSITPCTCGSSDLVLTADVADVPVRRRG
 DSRGSLSPRVSYLVKLLCPGSLGPHVGI FRAAVCTRGVAKADVPPVSEMETT
 TRSPVPTDSSPPAVPQFQVAHLHAPTGSGSTKVPAYAAQGYKVLVNPVSAATL
 GFGYMSKAGVGNENIKRTGRTITGAPITTSYTKRFLADGGCGGADYIMCDEPGK
 IDTSILGIVDQAEAGKRLGKDELAALGSLGNVAYITRYGLDVSVIPGSDVVVVA
 AIPRAIKGGRRHIFCHSKKDELAALGSLGNVAYITRYGLDVSVIPGSDVVVVA
 TDALMTGFDGLDVIDCNCVTQTFDFSLDPTFTTETTPQDRAVSQRGRTRG
 RRGIVFVTPGEPGSDSSVLCYDAGCAWYELTAEYTVARAVLNTPLPVCO
 DHEFWESVFTGLTHIDAHFSLQKQAGNLPYVAYQATVCARAAQPPPSDWQWKC
 LIRLKPGLHPTLLYRLGAVONEVILTHPIKYIMACMSADLEVTWVTLVGGVLA
 ALAYCLTGSVIVVGRILGSRPAVDPREVLYREDFEMERCASHLPIEGOMQAL
 QFKKALGELLQASQAABAAAPVPSKRALEAFWAKMNPISIGIYLAGLSTLPGN
 PAISLAFATASITPQTQSLTILGNGVAAQALAPPASAFASPGALVGVVCAA
 GLKVLVDIAGAGVAGALVAFVMSGEMPSDEDLVNLPAILSPGALVGVVCAA
 ILRHVGEAGVAGVOMNLLIAPASNGHVSPTHYPBESDAVVRVTLISLITOLLK
 RLHWINDECSYPCSGSLRDVMDVICTVLDFTKMLQSKLLPLRPVPPFLSCQGYK
 GWRGDSGIMHTTICPGCAQITGVRKNGSMRIVGPRICSNTHGTFVFNATGFCPTSP
 APNYSKALMRVAAEYVEVTRVGDVHYVTGTMTDNVKPCQCPVPEFFTEVDGVRLR
 YAPACKLLREBEVTFQVGLNQLVGSQPCPEPDAVLTSMITDPSHTATKRLR
 ARGSPSLASSASOLSAPSLKCTTRHDSPADLI EANLLRMDEMGNITRVSEEN
 KVVILDSPEPRAREDEVSAAEVLAKSRKFPALPIWARPDPYNPDLLESMDPDY
 IPIVHGCPLPTKAPPIPPRRKRTVLTSTVSALAEALATKTFGSESAVDSGV
 ATAPDQFDAGDSGYSYSPPLSGEPDPLSDGWSVSESEASEDVCCSMS
 DPPOVEDLEILTSSNVSAHDSGRVYLTTRDPTPLARAWETARPSPVNSWL
 GNIIMYPTLWARMLTTHFFSILLAQELKALDCQYGCYSTIEPDLQIOLRLH
 GLSAFSLHSYSGPNEINVASCLRGVPLRVLRARSVRKILSQCGRATCGKYL
 FNMAVTKLKLTPIPAAQSQDLNFWFVAGYSGDIIYHLSRARPMFWMCLLLLSVGV
 GYLLPNR"

ORIGIN

Query Match		90.8%;	Score 436;	DB 14;	Length 9379;
Best Local Similarity		94.7%;	Pred. No. 1.6e-101;		
Matches 451;		Conservative 0;	Mismatches 25;	Indels 0;	Gaps 0;
Qy	2	TGTCGGTGTCTTTCTATCTTCCTCTTGGCCCTCTGTCCTCTGACCACTACACAG	61		
Db	835	TGCCGGTGTCTTTCTATCTTCTCTTGGCTCTGTCCTCTGTCCTGTTTGACCACTCCAG	894		
Qy	62	CTTCGGCTATGAGTGGCGAACGTCGCGGGTGTACATGTACAGAACGACTGCTCA	121		
Db	895	CTTCGGCTATGAGTGGCGAACGTCGCGGGTGTACATGTACAGAACGACTGCTCA	954		
Qy	122	ACTCAAGCATAGTATGAGGAGCGGACATGATCATGCACACCCCGGTCGTCGCT	181		
Db	955	ACTCAAGCATAGTATGAGGAGCGGACATGATCATGCACACCCCGGTCGTCGCT	1014		
Qy	182	CGCTTCGGGAGCGCACTCTCCCTGCTGGTGGCGCTCACTCCACGCTCGCGGCA	241		
Db	1015	CGCTTCGGGAGCACTCTCCCTGCTGGTGGCGCTCACTCCACGCTCGCGGCA	1074		
Qy	242	GGAACGCGAGCGTCCCAACAAGCAATACAGACCAAGTCAATTTGCTGCTGGGCTG	301		
Db	1075	GGAACGCGAGCGTCCCAACAAGCAATACAGACCAAGTCAATTTGCTGCTGGGCTG	1134		
Qy	302	CTGCTTTCTGTCGCTATGAGTGGGGATCTCTGGGATCTCTTTTCTGTTTCC	361		
Db	1135	CTGCTTTCTGTCGCTATGAGTGGGGGACCTCTGCGGATCTGTTTCTGCTCTCC	1194		
Qy	362	AGCTGTTCACTCTCTCACTCGCGGCAACAAAGTACAGGACTGCAACTGCTCAATCT	421		
Db	1195	AGCTGTTCACTCTCTCGGCTCGCGGCAATGAGACGCTACAGACTGCAATGCTCAATCT	1254		
Qy	422	ATCCCGGCAATGATACAGTCAACGATGGCTTGGGATATGATGATGAACCTGGTAA	477		
Db	1255	ATCCCGGCAATGATACAGTCAACGATGGCTTGGGATATGATGATGAACCTGGTAA	1310		

RESULT 12 HPC5TRJ4 LOCUS DEFINITION ACCESSION VERSION KEYWORDS SOURCE ORGANISM

REFERENCE AUTHORS

TITLE JOURNAL MEDLINE PUBMED

REFERENCE AUTHORS

TITLE JOURNAL MEDLINE PUBMED

REFERENCE AUTHORS

TITLE JOURNAL MEDLINE PUBMED

REFERENCE AUTHORS

TITLE JOURNAL MEDLINE PUBMED

REFERENCE AUTHORS

TITLE JOURNAL MEDLINE PUBMED

REFERENCE AUTHORS

TITLE JOURNAL MEDLINE PUBMED

REFERENCE AUTHORS

TITLE JOURNAL MEDLINE PUBMED

REFERENCE AUTHORS

TITLE JOURNAL MEDLINE PUBMED

REFERENCE AUTHORS

TITLE JOURNAL MEDLINE PUBMED

HPC5TRJ4
 Hepatitis C virus gene for structural protein, partial cds,
 isolate:HC-J4.
 D00832
 D00832.1 GI:221513
 structural protein.
 Hepatitis C virus
 Hepatitis C virus
 Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 Hepacivirus.

1 (bases 1 to 1880)
 Okamoto, H., Okada, S., Sugiyama, Y., Yotsumoto, S., Tanaka, T.,
 Yoshizawa, H., Tada, F., Miyakawa, Y. and Mayumi, M.
 The 5'-terminal sequence of the hepatitis C virus genome
 Jpn. J. Exp. Med. 60 (3), 167-177 (1990)

2 (sites)
 Okamoto, H., Okada, S., Sugiyama, Y., Kurai, K., Iizuka, H., Machida, A.,
 Miyakawa, Y. and Mayumi, M.
 Nucleotide sequence of the genomic RNA of hepatitis C virus
 isolated from a human carrier: comparison with reported isolates
 for conserved and divergent regions
 J. Gen. Virol. 72 (Pt 11), 2697-2704 (1991)

3 (sites)
 Mink, M.A., Benichou, S., Madaule, P., Tiollais, P., Prince, A.M. and
 Inchausti, G.
 Characterization and mapping of a B-cell immunogenic domain in
 hepatitis C virus E2 glycoprotein using a yeast peptide library
 Virology 200 (1), 246-255 (1994)

4 (sites)
 Hotta, H., Doi, H., Hayaishi, T., Purwanta, M., Soemarto, W.,
 Mizokami, M., Ohba, K. and Homma, M.
 Analysis of the core and E1 envelope region sequences of a novel
 variant of hepatitis C virus obtained in Indonesia
 Arch. Virol. 136 (1-2), 53-62 (1994)

These data kindly submitted in computer readable form by: Hiroaki
 Okamoto
 Immunology Division, Jichi Medical School
 Kawachi-gun, Tochigi-ken 329-04
 Japan
 Phone: 0285-44-2111 x3334
 Fax: 0285-44-1557.

Location/Qualifiers
 1. 1880
 /organism="Hepatitis C virus"
 /mol_type="genomic RNA"
 /isolate="HC-J4"
 /db_xref="taxon:1103"
 /note="82 bp upstream of StyI site"
 342. .>1880
 /note="putative"
 /codon_start=1
 /product="structural protein"
 /protein_id="BAA00706.1"

/db_xref="GI:221514"
 /translation="MSTNPQRKTRNRRPQVKGPGQIVGVVLLPRRGRL
 GVRAKTSERSQPRGQPIKARRPGRWAQPGYPMPLXNGRGLWAGWLLSPRG
 SRVSGPTDPRRSNKLKVIDLTGCGFADLMGYIPLVGAPLGGARALAHVRLD
 GPNVATNLPGCSFIFLLALSLCITIPASAYEVNRVNSGIYHTVNDNSSIIVYRAAD
 VMHTPGCVPCVREDNSRCWVLTPLAARNASVPTTIRRHVDLLVGAAPFCAMV
 VDLCGSLVQLPFTFPRHETVQDCNCSYPGHLSGHRMAMMMNSPTTALV
 SOLLRIPQAVVDMVAGAHVGLAGLYSMVGNNAKVLIVALLFAGVDGTYTSGAA
 SHVTSTLASLSPGASQRIQLVNTNGSWHINETALNDSLHTFLAALFYTHRFNS

[illegible]

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2004, 07:39:53 ; Search time 180.6 Seconds
(without alignments)
11290.892 Million cell updates/sec

Title: US-09-899-303A-11
Perfect score: 480
Sequence: 1 ATGTCGGTGTCTCTTCTC.....TGATGATGAACGTGGTAATAG 480

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*

- 1: Geneseqn1980a:*
- 2: Geneseqn1990a:*
- 3: Geneseqn2000a:*
- 4: Geneseqn2001a:*
- 5: Geneseqn2001b:*
- 6: Geneseqn2002a:*
- 7: Geneseqn2003a:*
- 8: Geneseqn2003b:*
- 9: Geneseqn2003cs:*
- 10: Geneseqn2004a:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	480	100.0	480	2	AAT12708 HCV E1 co
2	480	100.0	480	6	AAL48917 Hepatitis
3	480	100.0	480	9	ADD55519 Hepatitis
4	474.2	98.8	483	2	AAT12707 HCV E1 co
5	474.2	98.8	483	6	AAL48916 Hepatitis
6	474.2	98.8	483	9	ADD55517 Hepatitis
7	432.8	90.2	1880	2	AQ24467 NANB hepa
8	432.8	90.2	2187	2	ABA03491 Cuticle p
9	432.8	90.2	2540	2	AQ43889 NANB hepa
10	432.8	90.2	2540	2	AQ63753 NANBHV ge
11	431.2	89.8	1863	2	AQ15363 Fragment
12	429.6	89.5	2540	2	AQ29628 Hepatitis
13	428	89.2	9595	2	AAX24843 Infection
14	428	89.2	9595	4	AAT23492 Infection
15	428	89.2	9595	4	AAC86939 Nucleotid
16	428	89.2	9599	2	AAX24833 Infection
17	426.4	88.8	577	2	AQ35081 HCV envel
18	426.4	88.8	2187	2	ABA03492 Cuticle p
19	426.2	88.8	633	2	AAT12706 HCV E1 co
20	426.2	88.8	633	6	AAL48915 Hepatitis
21	426.2	88.8	633	9	ADD55515 Hepatitis
22	424.6	88.5	567	2	AQ27160 NANB hepa
23	423.2	88.2	577	2	AQ35085 HCV envel

24	423.2	88.2	580	2	AAQ11076 Fragment
25	423.2	88.2	580	2	Aaz07647 HCV J1 E
26	423.2	88.2	1562	2	AAV60672 Fragment
27	423.2	88.2	1953	7	AAL55222 Plasmid p
28	423.2	88.2	2829	2	AAV60673 Fragment
29	422.8	88.1	580	2	AaQ79750 Hepatitis
30	422.8	88.1	580	2	AAX00401 Hepatitis
31	422.8	88.1	580	2	AAX26733 Consensus
32	422.8	88.1	580	2	AAX26728 Consensus
33	422.8	88.1	1249	2	AaQ79772 Hepatitis
34	422.8	88.1	1249	2	AAX26739 Consensus
35	422.8	88.1	1270	2	AAV60668 Fragment
36	422.8	88.1	3401	2	AAQ64069 Non-A, no
37	422.8	88.1	3401	2	AAT30387 5'UTR/COR
38	422.8	88.1	3461	2	AQ64068 Non-A, no
39	422.8	88.1	3461	2	AAT30386 5'UTR/COR
40	422.2	88.0	642	2	AAT12704 HCV E1 co
41	422.2	88.0	642	6	AAL48913 Hepatitis
42	422.2	88.0	642	9	ADD55634 Hepatitis
43	421.6	87.8	1682	2	AQ27159 NANB hepa
44	421.2	87.8	577	2	AAQ35086 HCV envel
45	421.2	87.8	580	2	AaQ79758 Hepatitis

ALIGNMENTS

RESULT 1

AAT12708
ID AAT12708 standard; DNA; 480 BP.

AC AAT12708;
DT 23-SEP-1996 (first entry)

XX HCV E1 construct HCC113A.

XX HCV; E1; E2; disulphide bond cleavage; envelope protein; vaccine; human;

KW serotype; reversed phase hybridisation assay; genotype; antigen; sera;

KW ss.

XX Hepatitis C virus.

XX WO9604385-A2.

PD 15-FEB-1996.

PP 31-JUL-1995; 95WO-EP003031.

XX 29-JUL-1994; 94EP-00870132.

PR (INNO-) INNOGENETICS NV.

XX Maertens G, Bosman F, De Martynoff G, Buyse M;

XX WPI; 1996-129401/13.

PT Purifying recombinant hepatitis C virus (HCV) E1 and E2 envelope proteins
- in presence of di: sulphide bond cleavage agent, to produce proteins
suitable for direct use in vaccines or diagnostic assays of HCV.
Claim 23; Fig 21; 146pp; English.

CC AAT12704-T12709 and AAT12961-T12974 represent hepatitis C virus (HCV) E1
and E2 protein coding sequence constructs. These sequences are included
in vectors for the production of recombinant E1, E2, and E1/E2 proteins.
The recombinant proteins can then be isolated using a method of the
invention. In the method, the envelope proteins are purified by carrying
out a disulphide bond cleavage, or a reduction step with a disulphide
bond cleavage agent, after lysis of recombinant host cells. The
constructs containing the purified HCV envelope proteins can be used for
vaccinating humans against HCV, for in vitro detection of HCV antibodies
in a sample, and in a serotyping assay for detecting one or more

Db	181	TGCGTTGGGAGGGCAA	CTCTCTCCGTTGCTGGGTGGCGCTCACTCCACGCTCGCGGCT	240
Qy	241	AGGAAAGCGCAGCGTCC	CACAAACGACAATATACGACGCCACGTCGATTTGCTCGTTGGGGCT	300
Db	241	AGGAAAGCGCAGCGTCC	CACAAACGACAATATACGACGCCACGTCGATTTGCTCGTTGGGGCT	300
Qy	301	GCTGCTTCTGCTCCGCTAT	GTACGTGGGGGATCTCTGCGGATCTGTTTTCCTTGTGTTCC	360
Db	301	GCTGCTTCTGCTCCGCTAT	GTACGTGGGGGATCTCTGCGGATCTGTTTTCCTTGTGTTCC	360
Qy	361	CAGCTGTTTCACCTTCT	CACCTCGCCGGGANTAAACAGTACAGGACTGCAACTGCTCAATC	420
Db	361	CAGCTGTTTCACCTTCT	CACCTCGCCGGGANTAAACAGTACAGGACTGCAACTGCTCAATC	420
Qy	421	TATCCGGCCATGTATAC	AGGTACCGCATGGCTTGGGATATGATGATGAACCTGGTAATA	479
Db	421	TATCCGGCCATGTATAC	AGGTACCGCATGGCTTGGGATATGATGATGAACCTGGTCCTA	479

RESULT 5

AAL48916		
ID	AAL48916	standard; DNA; 483 BP.
XX		
AC	AAL48916;	
XX		
DT	24-OCT-2002	(first entry)
XX		
DE	Hepatitis C virus clone HCC12A E1	protein coding sequence.
XX		
KW	Hepatitis C virus; HCV; E1 protein;	E2 protein; infection; gene;
XW	viruside; immunostimulant; vaccine;	ds.
XX		
OS	Hepatitis C virus.	
XX		
PN	WO200255548-A2.	
XX		
PD	18-JUL-2002.	
XX		
PF	11-JAN-2002; 2002WO-BF000219.	
XX		
PR	11-JAN-2001; 2001US-0260669P.	
PR	30-AUG-2001; 2001US-0315768P.	
XX		
PA	(INNO-) INNOGENETICS NV.	

Best Local Similarity 99.4%; Pred. No. 1.5e-128;		
Matches 476;	Conservative 0;	Mismatches 3; Indels 0; Gaps 0;
QY	1	ATGTCGCGGTGTCTTTCTCTATCTTCCTCTTGGCCCTGCTGCTCCTGCTGACCATACCA 60
Db	1	ATGCCCGGTGTCTTTCTCTATCTTCCTCTTGGCCCTGCTGCTCCTGCTGACCATACCA 60
QY	61	GCTTCCGCTTATGAAGTGCACACGNTGTCGGGGTGTAACATGTACAGAAAGCATGCTCC 120
Db	61	GCTTCCGCTTATGAAGTGCACACGNTGTCGGGGTGTAACATGTACAGAAAGCATGCTCC 120
QY	121	AACTCACGCAATAGTGATGAGGACGCGGACATGATCATGACACACCCCGGGTGGCTGCC 180
Db	121	AACTCACGCAATAGTGATGAGGACGCGGACATGATCATGACACACCCCGGGTGGCTGCC 180
QY	181	TGCGTTTCGGGAGGGCAACTCTCTCCGTTTGCTGGGTGGCGCTCACTCCACGCTCGCGGCC 240
Db	181	TGCGTTTCGGGAGGGCAACTCTCTCCGTTTGCTGGGTGGCGCTCACTCCACGCTCGCGGCC 240
QY	241	AGGAACGCCAGCGTCCCACAAACGACAAATACGACGCCACGTCGATTTGCTCGTTGGGGCT 300
Db	241	AGGAACGCCAGCGTCCCACAAACGACAAATACGACGCCACGTCGATTTGCTCGTTGGGGCT 300
QY	301	GCTGCTTTCTGTTCCGCTATGTAGTGGGGGATCTCTGGCGGATCTGTTTCCCTGTTTCC 360
Db	301	GCTGCTTTCTGTTCCGCTATGTAGTGGGGGATCTCTGGCGGATCTGTTTCCCTGTTTCC 360
QY	361	CAGCTGTTTCACTTCTCACTTCGCGCGGCATCAAAACAGTACAGGACTGCAACTGCTCAATC 420
Db	361	CAGCTGTTTCACTTCTCACTTCGCGCGGCATCAAAACAGTACAGGACTGCAACTGCTCAATC 420
QY	421	TATCCCGGCCATGTATCAGGTCAACCGCATGGCTTTGGGATATGATGATGAATGTGTAATA 479
Db	421	TATCCCGGCCATGTATCAGGTCAACCGCATGGCTTTGGGATATGATGATGAATGTGTAATA 479

RESULT 6

RESOUR 6	
ADD55517	
XX	ADD55517 standard; DNA; 483 BP.
XX	
AC	ADD55517;
XX	
DT	15-JAN-2004 (first entry)
XX	
DE	Hepatitis C virus E1 protein coding sequence #4.
XX	
KW	Hepatitis C virus; HCV; vaccine; liver disease; E1 protein; E2 protein;
KW	liver fibrosis; ds; gene.
XX	
OS	Hepatitis C virus.
XX	
FN	WO2003051912-A2.
XX	
PD	26-JUN-2003.
XX	
PF	18-DEC-2002; 2002WO-EP014480.
XX	
PR	18-DEC-2001; 2001US-00020510.
PR	16-OCT-2002; 2002US-0418358P.
XX	
PA	(INNO-) INNOGENETICS NV.
XX	
PI	Maertens G, Depla E, Bosman F;
XX	
DR	WPI; 2003-541632/51.
DR	P-PSDE; ADD55518.
XX	
PT	New hepatitis C virus (HCV) vaccine composition, useful for reducing
PT	liver disease, e.g., liver fibrosis in a chronic HCV-infected mammal.
XX	
PS	Example 1; SEQ ID NO 9; 271pp; English.
XX	
CC	The invention comprises an Hepatitis C virus (HCV) vaccine for reducing

QY	302	CTGCTTTTCTGTTCGGCTATGTACAGTGCGGGGATCTCTCGGATCTGTTCCTTGTTTTCC	367
Db	1147	CTGCCTTTCTGCTCCGCTATGTACAGTGCGGGGATCTCTCGGATCTGTTCCTTGCTCTCC	1206
QY	362	AGCTGTTCACCTTCTCACCCTCGCGGCATCAAAACAGTACAGACTGCCAAGCTGCAACTGCTCAATCT	421
Db	1207	AGCTGTTCACCTTCTCGGCTCGCGGCATGAGACAGTGCAGAGTGCMACTGCTCAATCT	1266
QY	422	ATCCCGGCATGTATCAGGTCACCGCATGGCTTGGGATATGATGATGAAGCTGGTAA	477
Db	1267	ATCCCGGCATTATCAGGTCACCGCATGGCTTGGGATATGATGATGAAGCTGGTCA	1322
 RESULT 10 AAQ63753 standard; cDNA to mRNA; 2540 BP.			
ID	AAQ63753		
XX	AAQ63753;		
XX	30-JAN-1995 (first entry)		
DT	NANBHV genomic fragment #2.		
DE	Polymerase chain reaction; PCR; primer; amplify; detection; NANBHV;		
KW	non-A, non-B hepatitis virus; 5'-terminal region; core protein; ss.		
KW	Synthetic.		
OS			
XX	JP06125777-A.		
FN	10-MAY-1994.		
PD			
PF	20-JUN-1991; 91JP-00247120.		
XX			
PR	20-JUN-1991; 91JP-00247120.		
XX	(NAKA/) NAKAMURA T.		
PA	WPI; 1994-187937/23.		
DR			
PT	Oligonucleotide primer pairs specific for non-A, non-B hepatitis virus -		
PT	used for high sensitivity detection of non-A non-B (NANB) hepatitis		
PT	virus.		
XX			
PS	Disclosure; Page 24-25; 25pp; Japanese.		
XX			
CC	The sequences given in AAQ63752-53 represent fragments of the non-A, non-		
CC	B hepatitis virus (NANBHV) genome. These fragments were amplified using		
CC	the primers given in AAQ63732-51. These primers were used in the		
CC	detection of NANBH. The primers are based on the 5'-terminal region and		
CC	the core protein coding region. The method allows highly sensitive		
CC	detection of NANBH		
XX			
SQ	Sequence 2540 BP; 470 A; 775 C; 742 G; 553 T; 0 U; 0 Other;		
 Query Match 90.2%; Score 432.8; DB 2; Length 2540; Best Local Similarity 94.3%; Pred. No. 3.5e-116; Matches 449; Conservative 0; Mismatches 27; Indels 0; Gaps 0;			
QY	2	TGTCGGGTTCCTTCTTCTATCTTCCTCTGGCCCTGCTGTCTGTCTGACCATACCAG	61
Db	847	TGCCGGTTCCTTCTTCTATCTTCTTCTTGGCTTGTCTGTGACCATCCAG	906
QY	62	CTTCCGCTTATGAAGTCGGCAAAGTCGCCGGGTGTACCATGTACAAGAGACTGTCCA	121
Db	907	CTTCCGCTTATGAAGTCGGCAAAGTCGCCGGATATAACATGTACAGAACGACTGTCCA	966
QY	122	ACTCMAGCATGTATGAGGACGCGACATGATCATGCACACCCCGGTGGTGCCT	181
Db	967	ACTCMAGCATGTATGAGGACGCGACATGATCATGATCTCCCGGTGGTGCCT	1026
QY	182	GCGTTCGGGAGGGCAACTCTCTCCGTTGCTGGGTGGCGCTCACTCCGACGCTCGCGCCA	241

[illegible]


```

PD XX 28-JAN-1999.
PF XX 16-JUL-1998; 98WO-US014688.
PR XX 18-JUL-1997; 97US-0053062P.
PR XX 27-JAN-1998; 98US-00014416.
PR XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA XX Yanagi M, Bukh J, Emerson SU, Purcell RH;
PI XX WPI; 1999-132252/11.
XX XX P-PSDB; AAW98022.
XX XX
PT XX New isolated hepatitis C virus nucleic acids - used to develop products
PT XX for the diagnosis, prevention and treatment of HCV infections and for
PT XX developing screening assays.
XX XX
PS XX Claim 3; Fig 14A-F; 126pp; English.
XX XX
CC XX The present sequence comprises the nucleic acid sequence of the genome of
CC XX infectious hepatitis C virus (HCV) genotype 1b strain HC-J4 (ATCC 209596)
CC XX that is capable of expressing this virus when transfected into cells. HC-
CC XX J4 was obtained from acute phase plasma of a chimpanzee experimentally
CC XX infected with serum containing HC-J4/91. The claimed infectious nucleic
CC XX acid sequence can be used to produce chimeric genomes (see AAX24833)
CC XX consisting of the open reading frames of infectious nucleic acid
CC XX sequences of other genotypes (including genotypes 1-6) and subtypes (such
CC XX as 1b, 2a, 2b, 2c, 3a, 4a-f, 5a and 6a) of HCV. The invention also
CC XX relates to the introduction of mutations or deletions into infectious
CC XX nucleic acid sequences in order to produce an attenuated HCV virus
CC XX suitable for vaccine development. Infectious nucleic acid sequences can
CC XX also be used to produce attenuated virus via passage in vitro or in vivo
CC XX of the viruses produced by transfection of a host cell with the
CC XX infectious nucleic acid sequence. Vaccines comprising one or more
CC XX polypeptides made from the infectious nucleic acid sequence are used to
CC XX immunise mammals, especially humans, against hepatitis C. The nucleic
CC XX acid sequences can also be used to induce protective immunity against the
CC XX virus. The nucleic acid sequences or their encoded proteases (e.g. NS3
CC XX protease) can additionally be used to develop screening assays to
CC XX identify antiviral agents for HCV
XX XX
SQ XX Sequence 9595 BP; 1934 A; 2842 C; 2698 G; 2121 T; 0 U; 0 Other;
Query Match 89.2%; Score 428; DB 2; Length 9595;
Best Local Similarity 93.7%; Pred. No. 1.5e-114;
Matches 446; Conservative 0; Mismatches 30; Indels 0; Gaps 0;
Qy 2 TGTCGGTGTGCTTTCTCTATCTCTCTTGGCCCTGCTGCTCTGTGTGACCATACAG 61
Db 847 TGCCCGGTTGCTCTTTCTCTATCTCTCTTGGCTGCTGCTCTGTGTGACCATCCAG 906
Qy 62 CTTCCGGTTATGAAGTGGCAACGTGTCCGGGTGTACCATGTCAAGAACGACTGCTCA 121
Db 907 CTTCCGGTTATGAAGTGGCAACGTGTCCGGGTGTACCATGTCAAGAACGACTGCTCA 966
Qy 122 ACTCAAGCATAGTGTATGAGCGAGGACATGATCATGTCACACCCCGGGTGGTGCCT 181
Db 967 ACTCAAGCATAGTGTATGAGCGAGGACATGATCATGTCACACCCCGGGTGGTGCCT 1026
Qy 182 CGGTTCCGGAGGGCAACTCTCTCCGTTGCTGGGTGGGCTCACTCCCAAGCTCGCGGCA 241
Db 1027 GTGTTCCAGAGGGTAAACAGTCTCCGTTGCTGGGTAGGCTCACTCCCAAGCTCGCGGCA 1086
Qy 242 GGAAAGCAGCGTCCCAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 301
Db 1087 GGAAAGCAGCGTCCCAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 1146
Qy 302 CTGCTTTCTGTCGCTATGATGCTGGGGGATCTCTGGGATCTGTTTCTGCTGTTCCC 361
Db 1147 CTGCTTTCTGTCGCTATGATGCTGGGGGATCTCTGGGATCTGTTTCTGCTGTTCCC 1206
Qy 362 AGCTGTTTCACTTCTACCTCGCCGGCATCAACAGTACAGGACTGCAACTGCTCAATCT 421

```

```

Db 1207 AGCTGTTTCACTTCTGCTCGCTCGGCATGACACAGTGCAGACTGCATCTCAATCT 1266
Qy 422 ATCCCGGGCATGTATCAGGTCAACCGCATGGCTTGGGATATGATGATGAACCTGTAA 477
Db 1267 ATCCCGGGCATGTATCAGGTCAACCGCATGGCTTGGGATATGATGATGAACCTGTAA 1322
RESULT 14
AAP23492
ID AAP23492 standard; DNA; 9595 BP.
XX AC AAP23492;
XX XX
DT 21-MAR-2001 (first entry)
DE Infectious Hepatitis C virus 1b genotype.
XX KW GBV-B; hepatitis C virus; HCV; vaccine; ds.
XX OS Hepatitis C virus.
XX PN WO2000075337-A1.
XX PD 14-DEC-2000.
XX PF 02-JUN-2000; 2000WO-US015293.
XX PR 04-JUN-1999; 99US-0137694P.
XX XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX PI Bukh J, Yanagi M, Emerson SU, Purcell RH;
XX DR WPI; 2001-091214/10.
XX PT New infectious nucleic acids of the GB virus-B clone, useful for
XX PT indirectly studying the molecular properties of hepatitis C virus (HCV)
XX PT and in developing vaccines and therapeutics for HCV.
XX PS Disclosure; Fig 7; 96pp; English.
XX XX
CC XX The present invention relates to GB virus-B. The nucleic acid molecules
CC XX of the invention are useful for indirectly studying the molecular
CC XX properties of hepatitis C virus (HCV). The infectious nucleic acid
CC XX sequence of the GB virus-B clone and the HCV/GBV-B chimeras may be used
CC XX in the development of vaccines and therapeutics for HCV
XX XX
SQ XX Sequence 9595 BP; 1934 A; 2843 C; 2697 G; 2121 T; 0 U; 0 Other;
Query Match 89.2%; Score 428; DB 4; Length 9595;
Best Local Similarity 93.7%; Pred. No. 1.5e-114;
Matches 446; Conservative 0; Mismatches 30; Indels 0; Gaps 0;
Qy 2 TGTCGGTGTGCTTTCTCTATCTCTCTTGGCCCTGCTGCTCTGTGTGACCATACAG 61
Db 847 TGCCCGGTTGCTCTTTCTCTATCTCTCTTGGCTGCTGCTCTGTGTGACCATCCAG 906
Qy 62 CTTCCGGTTATGAAGTGGCAACGTGTCCGGGTGTACCATGTCAAGAACGACTGCTCA 121
Db 907 CTTCCGGTTATGAAGTGGCAACGTGTCCGGGTGTACCATGTCAAGAACGACTGCTCA 966
Qy 122 ACTCAAGCATAGTGTATGAGCGAGGACATGATCATGTCACACCCCGGGTGGTGCCT 181
Db 967 ACTCAAGCATAGTGTATGAGCGAGGACATGATCATGTCACACCCCGGGTGGTGCCT 1026
Qy 182 CGGTTCCGGAGGGCAACTCTCTCCGTTGCTGGGTGGGCTCACTCCCAAGCTCGCGGCA 241
Db 1027 GTGTTCCAGAGGGTAAACAGTCTCCGTTGCTGGGTAGGCTCACTCCCAAGCTCGCGGCA 1086
Qy 242 GGAAAGCAGCGTCCCAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 301
Db 1087 GGAAAGCAGCGTCCCAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 1146

```


GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: June 16, 2004, 09:12:53 ; Search time 1290.4 Seconds
(without alignments)
11108.065 Million cell updates/sec

Title: US-09-899-303A-11
Perfect score: 480
Sequence: 1 ATGTCGGTTCCTTTCTC.....TGATGATGACGTGTAATAG 480

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :	EST:
1: em_estba:*	
2: em_esthum:*	
3: em_estin:*	
4: em_estmu:*	
5: em_estov:*	
6: em_estpl:*	
7: em_estro:*	
8: em_hcc:*	
9: gb_est1:*	
10: gb_est2:*	
11: gb_hcc:*	
12: gb_est3:*	
13: gb_est4:*	
14: gb_est5:*	
15: em_estfun:*	
16: em_estom:*	
17: em_gss_hum:*	
18: em_gss_inv:*	
19: em_gss_pin:*	
20: em_gss_vrt:*	
21: em_gss_fun:*	
22: em_gss_mam:*	
23: em_gss_mus:*	
24: em_gss_pro:*	
25: em_gss_rod:*	
26: em_gss_phg:*	
27: em_gss_vrl:*	
28: gb_gss1:*	
29: gb_gss2:*	

ALIGNMENTS

RESULT 1	AV755731/c	488 bp	mRNA	linear	EST 19-OCT-2000
LOCUS	AV755731	BM Homo sapiens	cDNA Clone BMAPKB03 5'		mRNA sequence.
DEFINITION	AV755731				
ACCESSION	AV755731				
VERSION	AV755731.1	GI:10913579			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 488)				
AUTHORS	Gu,J., Zhao,M., Huang,Q., Xu,X., Li,Y., Peng,Y., Song,H., Xiao,H., Gu,Y., Li,N., Qian,B., Liu,F., Qu,J., Gao,X., Cheng,Z., Xu,Z., Zeng,L., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Yang,Y., Gao,G., Wang,Z., Zhang,Q., Chen,S., Han,Z. and Chen,Z.				
TITLE	Homo sapiens cDNA BM clones				
JOURNAL	Unpublished (2000)				
COMMENT	Contact: Zeguang Han Chinese National Human Genome Center at Shanghai 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China Tel: 86-21-50801919(ex.45) Fax: 86-21-50801922 Email: hanzg@chgc.sh.cn This clone is available at CHGC in Shanghai.				

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
c 1	65.8	13.7	488	9	AV755731	AV755731
c 2	54.6	11.4	492	9	AV758366	AV758366
c 3	43.2	9.0	534	14	CF846043	CF846043
c 4	41.6	8.7	564	28	BZ645446	BZ645446

QY 299 CTGCTGCTTCTGTTCCGCT 318
||| | ||| |||
Db 86 GTGCTGCTGCTGCTGCT 67

RESULT 6	ACCESSION	ORGANISM
CG213867/c	VERSION	
LOCUS	KEYWORDS	
DEFINITION	SOURCE	

CG213867 794 bp DNA linear GSS 22-AUG-2003
CGVCR33TV ZM_0.7_1.5 KB Zea mays genomic clone ZMWBMA0506110,
genomic survey sequence.
CG213867
CG213867.1 GI:34113755
GSS.
Zea mays
Zea mays
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
Clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 794)

Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.

Consortium for Maize Genomics

Unpublished (2002)

Other GSSs: OGVCR53TH

Contact: Cathy Whitelaw

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843

Fax: 301-838-0208

Email: whitelaw@tigr.org

Seq primer: TP

Class: sheared ends.

Location/Qualifiers

1..794

/organism="Zea mays"

/mol_type="genomic DNA"

/strain="B73"

/db_xref="taxon:4577"

/clone="ZMMBma0506110"

/clone_lib="ZM 0.7 1.5 KB"

/note="vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN

Query Match 8.7%; Score 41.6; DB 29; Length 794;
Best Local Similarity 50.5%; Pred. No. 10;
Matches 101; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

QY 119 CCAACTCAAGCATAGTGTATGAGCAGCGGACATGATCATGCACACCCCGGGTGGTC 178
DB 397 CCAGCAGACGACACGTCGTGGTGGCGCCCAAGAGCGAGGTGCGGAAGCGGGCCGGAC 338
QY 179 CTGCGTTCGGGAGGCGCAACTCTCCGTTGCTGGGTGGCGCTCACTCCACGCTCGCGG 238
DB 337 CGTACCTCGGTGGCGAGCCTCCAGGAAGTGGTAGAGCCCTGCGCTGCTGTTGGCGG 278
QY 239 CCAGGACCGCAGGTCCTCCACAGCAGCAATFACAGCCGAGTGCATTTCTGTTGGGG 298
DB 277 ACAGGAAGCGCCAGCGTCTCCACGACGAGCACCACCGAGCGGTACCCCTTCTCGGTGGTG 218
QY 299 CTGCTGCTTTCTGTTCCGCT 318
DB 217 GTGGTGGTGGTGGTGGTCT 198

RESULT 7

CC634840/c

LOCUS

DEFINITION OGVCB40TV ZM 0.7 1.5 KB Zea mays genomic clone ZMMBma0500G08,
genomic survey sequence.

ACCESSION CC634840

VERSION CC634840.1

KEYWORDS GSS.

SOURCE Zea mays

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 995)

Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.

Consortium for Maize Genomics

Unpublished (2002)

Other GSSs: OGVCB40TH

Contact: Cathy Whitelaw

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA

FEATURES

source

Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TP
Class: sheared ends.
Location/Qualifiers
1..995
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBma0500G08"
/clone_lib="ZM 0.7 1.5 KB"
/note="vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN

Query Match 8.7%; Score 41.6; DB 29; Length 995;
Best Local Similarity 50.5%; Pred. No. 11;
Matches 101; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

QY 119 CCAACTCAAGCATAGTGTATGAGCAGCGGACATGATCATGCACACCCCGGGTGGTC 178
DB 864 CCAGCAGACGACACGTCGTGGTGGCGCCCAAGAGCGAGGTGCGGAAGCGGGCCGGAC 805
QY 179 CTGCGTTCGGGAGGCGCAACTCTCCGTTGCTGGGTGGCGCTCACTCCACGCTCGCGG 238
DB 804 CGTACCTCGGTGGCGAGCCTCCACGAAGTGGTAGAGCCCTGCGCTGCTGTTGGCGG 745
QY 239 CCAGGAACCGCAGCGTCCCAACAGCAATFACAGCCGAGTGCATTTGCTGTTGGGG 298
DB 744 ACAGGAAGCGCCAGCGTCTCCACGACGAGCACCACCGAGCGGTACCCCTTCTCGGTGGTG 685
QY 299 CTGCTGCTTTCTGTTCCGCT 318
DB 684 GTGGTGGTGGTGGTGGTCT 665

RESULT 8

CG373319/c

LOCUS

DEFINITION OGLBI44TV ZM 0.7 1.5 KB Zea mays genomic clone ZMMBma0723H15,
genomic survey sequence.

ACCESSION CG373319

VERSION CG373319.1

KEYWORDS GSS.

SOURCE Zea mays

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 761)

Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.

Consortium for Maize Genomics

Unpublished (2002)

Other GSSs: OGLBI44TH

Contact: Cathy Whitelaw

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843

Fax: 301-838-0208

Email: whitelaw@tigr.org

Seq primer: TP

Class: sheared ends.

Location/Qualifiers

1..761

/organism="Zea mays"

/mol_type="genomic DNA"

/strain="B73"

/db_xref="taxon:4577"

/clone="ZMMBma0723H15"

Db 376 C 376

Search completed: June 16, 2004, 13:27:14
Job time : 1291.4 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2004, 09:14:33 ; Search time 34.6 Seconds
(without alignments)
7698.741 Million cell updates/sec

Title: US-09-899-303A-11

Perfect score: 480

Sequence: 1 ATGTCGGTTCCTCTTCTC.....TGATGATGACTGGTAATAG 480

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA.*
1: /cgm2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgm2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgm2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgm2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgm2_6/ptodata/2/ina/ECTUS_COMB.seq.*
6: /cgm2_6/ptodata/2/ina/backfiles.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	480	100.0	480	3	US-08-612-973-11
2	480	100.0	480	3	US-08-927-597-11
3	474.2	98.8	483	3	US-08-612-973-9
4	474.2	98.8	483	3	US-08-927-597-9
5	431.2	89.8	1539	2	US-08-470-426B-17
6	431.2	89.8	1863	2	US-08-470-426B-14
7	428	89.2	9595	3	US-09-014-416-4
8	428	89.2	9599	3	US-09-014-416-6
9	426.2	88.8	633	3	US-08-612-973-7
10	426.2	88.8	633	3	US-08-927-597-7
11	423.4	88.2	501	2	US-08-483-695-30
12	423.4	88.2	501	2	US-07-965-285-30
13	423.4	88.2	501	2	US-08-487-231-30
14	423.4	88.2	501	3	US-09-201-912-30
15	422.8	88.1	9472	4	US-08-150-204E-36
16	422.2	88.0	642	3	US-08-612-973-3
17	422.2	88.0	642	3	US-08-927-597-3
18	421.2	87.8	795	3	US-08-612-973-5
19	421.2	87.8	795	3	US-08-927-597-5
20	421.2	87.8	2082	3	US-08-612-973-47
21	421.2	87.8	2082	3	US-08-927-597-47
22	421.2	87.8	2433	3	US-08-612-973-49
23	421.2	87.8	2433	3	US-08-927-597-49
24	419.6	87.4	1037	1	US-08-462-195-1
25	419.6	87.4	1037	2	US-08-636-883-1
26	419.6	87.4	1037	3	US-09-127-829-1
27	417.2	86.9	636	3	US-08-612-973-13

28	417.2	86.9	636	3	US-08-927-597-13	Sequence 13, Appl
29	416.8	86.8	742	1	US-08-081-072-18	Sequence 18, Appl
30	416.8	86.8	742	1	US-08-449-093A-18	Sequence 18, Appl
31	416.8	86.8	932	1	US-08-081-072-15	Sequence 15, Appl
32	416.8	86.8	932	1	US-08-449-093A-15	Sequence 15, Appl
33	415.4	86.5	2116	3	US-08-191-160-21	Sequence 21, Appl
34	413.8	86.2	501	2	US-08-483-695-28	Sequence 28, Appl
35	413.8	86.2	501	2	US-07-965-285-28	Sequence 28, Appl
36	413.8	86.2	501	3	US-08-487-231-28	Sequence 28, Appl
37	413.8	86.2	501	3	US-09-201-912-28	Sequence 28, Appl
38	413.6	86.2	11076	4	US-09-539-601-1	Sequence 1, Appl
39	413.6	86.2	11076	4	US-09-539-601-19	Sequence 19, Appl
40	413.6	86.2	11076	4	US-09-539-601-25	Sequence 25, Appl
41	413.6	86.2	11076	4	US-09-539-601-31	Sequence 31, Appl
42	405.2	84.4	1167	1	US-08-324-977-9	Sequence 9, Appl
43	405.2	84.4	1167	2	US-08-384-616-9	Sequence 9, Appl
44	405.2	84.4	1167	2	US-08-904-686A-9	Sequence 9, Appl
45	405.2	84.4	1167	3	US-09-315-850-9	Sequence 9, Appl

ALIGNMENTS

RESULT 1
US-08-612-973-11
; Sequence 11, Application US/08612973
; Patent No. 6150134
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,973
; FILING DATE: 11-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 480 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHEetical: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..477
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..474


```

CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/612,973
FILING DATE: 11-MAR-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1487-10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 483 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..480
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 1..477
US-08-612-973-9

```

Query Match	98.8%;	Score 474.2;	DB 3;	Length 483;
Best Local Similarity	99.4%;	Pred. No. 1.8e-125;		
Matches 476;	Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;
Qy	1	ATGTCGGGTTGCTCTTTCTCTATCTTCTCTTGGCCCTGCTGTGCTCTGTGACCATACCA	60	
Db	1	ATGCCCGGTTGCTCTTTCTCTATCTTCTCTTGGCCCTGCTGTGCTCTGTGACCATACCA	60	
Qy	61	GCTTTCGCTTATGAAGTGCACAGTGTGTCGGGGTGTAACCATGTCACGAAGACTGCTCC	120	
Db	61	GCTTTCGCTTATGAAGTGCACAGTGTGTCGGGGTGTAACCATGTCACGAAGACTGCTCC	120	
Qy	121	AACTCAAGCATAGTGTATGAGGCAGCGGACATGATCATGACACACCCCGGGTGCTGCGCC	180	
Db	121	AACTCAAGCATAGTGTATGAGGCAGCGGACATGATCATGACACACCCCGGGTGCTGCGCC	180	
Qy	181	TGGGTTCCGGAGGGCAACTCTCTCCGTTGCTGGGTGGGGTCACTCCACGCTCGCGGCC	240	
Db	181	TGGGTTCCGGAGGGCAACTCTCTCCGTTGCTGGGTGGGGTCACTCCACGCTCGCGGCC	240	
Qy	241	AGGAAGCAGAGCTCCCCCAACAGCAATACGACGCCACGTGATTCGTTGGGGCT	300	
Db	241	AGGAAGCAGAGCTCCCCCAACAGCAATACGACGCCACGTGATTCGTTGGGGCT	300	
Qy	301	GCTGCTTTCGTTTCGGCTATGTAGTGGGGGATCTCTGGGATCTGTTTCTTGTTTCC	360	
Db	301	GCTGCTTTCGTTTCGGCTATGTAGTGGGGGATCTCTGGGATCTGTTTCTTGTTTCC	360	
Qy	361	CAGCTGTTCACTTCTCACTCCGCGCATCAACAGTACAGGACTGCAACTGCTCAATC	420	
Db	361	CAGCTGTTCACTTCTCACTCCGCGCATCAACAGTACAGGACTGCAACTGCTCAATC	420	
Qy	421	TATCCCGGCCCATGTATCAGGTCAACCGATGGCTTGGGATATGATGATGAACTGGTAATA	479	
Db	421	TATCCCGGCCCATGTATCAGGTCAACCGATGGCTTGGGATATGATGATGAACTGGTCCTA	479	

```

RESULT 4
US-08-927-597-9
; Sequence 9, Application US/08927597
; Patent No. 6245503
; GENERAL INFORMATION:
; APPLICANT: MAERTENS, GEERT
; APPLICANT: BOSMAN, FONS
; APPLICANT: DE MARTYNOFF, GUY
; APPLICANT: BUYSSE, MARIE-ANGE
; TITLE OF INVENTION: PURIFIED HEPATITIS C VIRUS ENVELOPE
; TITLE OF INVENTION: PROTEINS FOR DIAGNOSTIC AND THERAPEUTIC USE
; NUMBER OF SEQUENCES: 111
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHUYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/927,597
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/612,973
; FILING DATE: 11-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 483 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..480
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 1..477
US-08-927-597-9

Query Match 98.8%; Score 474.2; DB 3; Length 483;
Best Local Similarity 99.4%; Pred. No. 1.8e-125;
Matches 476; Conservative 0; Mismatches 3; Indels 0;

QY 1 ATGTCGGGTTCCTTTCTCTATCTTCTCTGCGCCCTGCTCTCTGTCTGACCA
DB 1 ATGCCCGGTTCCTTTCTCTATCTTCTCTGCGCCCTGCTCTGTCTGACCA
QY 61 GCTTCGCGTTATGAAGTGCACCGTGTCCGGGGTTCATGTCAACGACGCT
DB 61 GCTTCGCGTTATGAAGTGCACCGTGTCCGGGGTTCATGTCAACGACGCT
QY 121 AACTCAACGATGCTATGAGGACCGGACATGATCATGACACACCCCGGGTGGC
DB 121 AACTCAACGATGCTATGAGGACCGGACATGATCATGACACACCCCGGGTGGC
QY 181 TCGCTTCCGGAGGGCACTCCTCCCGTTGCTGGGTGGCGCTCACTCCACGCTCTCG

```

Db 181 TGGCTTCGGAGGCAACTCTCCCGTGTGGGTGGGCTCACTCCCAAGCTCGCGCC 240
Qy 241 AGGAACGCCAGCGTCCCAACACGACAAATACGACGCGCACTCGATTTGCTGTGGGGCT 300
Db 241 AGGAACGCCAGCGTCCCAACACGACAAATACGACGCGCACTCGATTTGCTGTGGGGCT 300
Qy 301 GCTGCTTTCTGTCGCTATGTAGTACGTTGGGGGATCTCTGCGGATCTGTTTCTTTGTTCC 360
Db 301 GCTGCTTTCTGTCGCTATGTAGTACGTTGGGGGATCTCTGCGGATCTGTTTCTTTGTTCC 360
Qy 361 CAGCTGTTCACTTCTCACTCCCGCGCATCAACAGTACAGGACTCGAAGTCTCAATC 420
Db 361 CAGCTGTTCACTTCTCACTCCCGCGCATCAACAGTACAGGACTCGAAGTCTCAATC 420
Qy 421 TATCCCGGCCATGTATCAGGTACCGCATGCTTGGGATATGATGAACTGTAATA 479
Db 421 TATCCCGGCCATGTATCAGGTACCGCATGCTTGGGATATGATGAACTGTAATA 479

RESULT 5
US-08-470-426B-17
; Sequence 17, Application US/08470426B
; Patent No. 5856458
; GENERAL INFORMATION:
; APPLICANT: Okamoto, Hiroaki
; APPLICANT: Nakamura, Tetsuo
; TITLE OF INVENTION: OLIGONUCLEOTIDE PRIMERS, AND THEIR
; TITLE OF INVENTION: APPLICATION FOR HIGH-FIDELITY DETECTION OF NON-A, NON-B
; TITLE OF INVENTION: HEPATITIS VIRUS
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Beveridge, DeGrandi, Weillacher & Young,
; ADDRESSEE: L.L.P.
; STREET: 1850 M Street, N.W., Suite 800
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/470,426B
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-153402
; FILING DATE: 12-JUN-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Weillacher, Robert G.
; REGISTRATION NUMBER: 20,531
; REFERENCE/DOCKET NUMBER: 06/59-47083.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2811
; TELEFAX: (202) 659-1462
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1539 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
US-08-470-426B-17

Query Match 89.8%; Score 431.2; DB 2; Length 1539;
Best Local Similarity 94.1%; Pred. No. 4.2e-113;
Matches 448; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 2 TGGCGGTGCTCTTTCTCTATCTTCTTGGCTTTGCTCTGTTTGGACATCCAG 565

Db 565 TGGCGGTGCTCTTTCTCTATCTTCTTGGCTTTGCTCTGTTTGGACATCCAG 565
Qy 62 CTTCGCTTATGAAGTGGCAACGTTGTCGCGGGTGTACATGTACAGAACGACTCTCCA 121
Db 566 CTTCGCTTATGAAGTGGCAACGTTGTCGCGGGATATACATGTACAGAACGACTCTCCA 625
Qy 122 ACTCAAGCATATGTATGAGCGACGCGGACATGATGATCAACACCCCGGGTGGCGCT 181
Db 526 ACTCAAGCATATGTATGAGCGACGCGGACATGATGATCAACACCCCGGGTGGCGCT 685
Qy 182 GCGTTCGGGAGGGCAACTCTCCCGTGTGCTGGGTGGCGCTCACTCCACGCTCGCGGCA 241
Db 686 GCGTTCGGGAGGCAACTCTCCCGTGTGCTGGGTGGCGCTCACTCCACGCTCGCGGCA 745
Qy 242 GGAACGCCAGCGTCCCAACACGACAAATACGACGCGCATGCTGCTGTTGGGGCTG 301
Db 746 GGAATGCCAGGGTCCCACTACGACAAATACGACGCGCATGCTGCTGTTGGGGCTG 805
Qy 302 CTGCTTTCTGCTCGCTATGTACGTTGGGGGATCTCTGCGGATCTGTTTCTTGTTC 361
Db 806 CTGCTTTCTGCTCGCTATGTACGTTGGGGGATCTCTGCGGATCTGTTTCTTGTTC 865
Qy 362 AGCTGTTCACTTCTCACTCCCGCGCATCAACAGTACAGGACTCGAAGTCTCAATCT 421
Db 866 AGCTGTTCACTTCTCACTCCCGCGCATCAACAGTACAGGACTCGAAGTCTCAATCT 925
Qy 422 ATCCCGGCCATGTATCAGGTACCGCATGCTTGGGATATGATGAACTGTAATA 477
Db 926 ATCCCGGCCATGTATCAGGTACCGCATGCTTGGGATATGATGAACTGTAATA 981

RESULT 6
US-08-470-426B-14
; Sequence 14, Application US/08470426B
; Patent No. 5856458
; GENERAL INFORMATION:
; APPLICANT: Okamoto, Hiroaki
; APPLICANT: Nakamura, Tetsuo
; TITLE OF INVENTION: OLIGONUCLEOTIDE PRIMERS, AND THEIR
; TITLE OF INVENTION: APPLICATION FOR HIGH-FIDELITY DETECTION OF NON-A, NON-B
; TITLE OF INVENTION: HEPATITIS VIRUS
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Beveridge, DeGrandi, Weillacher & Young,
; ADDRESSEE: L.L.P.
; STREET: 1850 M Street, N.W., Suite 800
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/470,426B
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-153402
; FILING DATE: 12-JUN-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Weillacher, Robert G.
; REGISTRATION NUMBER: 20,531
; REFERENCE/DOCKET NUMBER: 06/59-47083.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2811
; TELEFAX: (202) 659-1462
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1863 base pairs
; TYPE: nucleic acid

[illegible]


```
; APPLICATION NUMBER: KR 91-13601
; FILING DATE: 6-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Shahan Islam, Esq.
; REGISTRATION NUMBER: 32,507
; REFERENCE/DOCKET NUMBER: 2695/FLK
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 940-8564
; TELEFAX: (212) 940-8564
; INFORMATION FOR SEQ ID NO: 96
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9472 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; FEATURE:
; OTHER INFORMATION: KHCV-LBC1, Fig. 2
; SEQUENCE DESCRIPTION: SEQ ID NO: 96
US-08-150-204E-96

Query Match      88.1%; Score 422.8; DB 4; Length 9472;
Best Local Similarity 93.2%; Pred. No. 1.8e-110;
Matches 442; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

Qy      2 TGTGGGTGCTCTTCTCTATCTTCTCTTGGCCCTGCTGCTCTGTACCATACCAG 61
Db      848 TGCCTGGTGGTCTCTTCTCTATCTTCTCTTGGCTCTGCTCTTGTGTTGACCCAG 907
Qy      62 CTTCGGCTTATGAGTGGCAGCGTCTCGGGGTGTACCATCTCAGCAACGACTGCTCCA 121
Db      908 TTTCGGCTTATGAGTGGTAACTCGGCTCGGGATGTACCATGTACGCAACGACTGCTCCA 967
Qy     122 ACTCAAGCATAGTGTATGAGGAGCAGCGGACATGATCATGACACACCCCGGGTGGTGGCT 181
Db     968 ACTCAAGCATTTGTATGAGGAGCGGACATGATCATGACACACTCCGGGTGGTGGCT 1027
Qy     182 GCGTTGGGAGGGCAACTCTCCCGTTGCTGGGTGGGCTCACTCCCAACGCTCGCGGCA 241
Db    1028 GCGTTGGGAGGCAAACTCTCCCGTTGCTGGGTGGCACTTACTCCCAACGCTCGCGGCA 1087
Qy     242 GGAACGCGCGTCCCAACAGCAATACGACGCCACGTCGATTTGCTGGGGCTG 301
Db    1088 GGAATGCGAGCTCCCACTACGACATTCGACGCCATGTGACTGCTCGTGGGGTAG 1147
Qy     302 CTGCTTTCTGTTCCGCTATGTACGTGGGGATCTCTGGGATCTGTTTCTTGTTCCTCC 361
Db    1148 CTGCTTTCTGTTCCGCTATGTACGTGGGGATCTCTGGGATCTGTTTCTTGTTCCTCC 1207
Qy     362 AGCTGTTACCTTCTACCTGCGCGGCATCAACAGTACAGGACTGCAACTGCTCAATCT 421
Db    1208 AGCTGTTACCTTTTCGCTCGCGGCATGAGACGGTACAGGACTGCAACTGCTCAATCT 1267
Qy     422 ATCCCGGCGGTATGATGAGTACCGCATGGCTTGGGATGATGATGATGATGATGAT 475
Db    1268 ATCCCGGCGGTATGAGTACCGCATGGCTTGGGATGATGATGATGATGATGATGAT 1321
```

Search completed: June 16, 2004, 13:30:24
Job time : 35.6 secs